EUROPEAN ECONOMY

EUROPEAN COMMISSION
DIRECTORATE-GENERAL FOR ECONOMIC AND FINANCIAL AFFAIRS

ECONOMIC PAPERS

ISSN 1725-3187
http://europa.eu.int/comm/economy_finance

N° 249 June 2006

The Stacked Leading Indicators Dynamic Factor Model: A Sensitivity Analysis of Forecast Accuracy using Bootstrapping

by
Daniel Grenouilleau
Directorate-General for Economic and Financial Affairs

Economic Papers are written by the Staff of the Directorate-General for Economic and Financial Affairs, or by experts working in association with them. The "Papers" are intended to increase awareness of the technical work being done by the staff and to seek comments and suggestions for further analyses. Views expressed represent exclusively the positions of the author and do not necessarily correspond to those of the European Commission. Comments and enquiries should be addressed to the:

European Commission
Directorate-General for Economic and Financial Affairs
Publications
BU1 - -1/13
B - 1049 Brussels, Belgium

ECFIN/REP 53331 - EN

ISBN 92-79-01190-1

KC-AI-06-249-EN-C

TABLE OF CONTENTS

THE STACKED LEADING INDICATORS DYNAMIC FACTOR MODEL: A SENSITIVITY ANALYSIS OF FORECAST ACCURACY USING BOOTSTRAPPING

Daniel GRENOUILLEAU*

Abstract:

The paper introduces an approximate dynamic factor model based on the extraction of principal components from a very large number of leading indicators stacked at various lags. The model is designed to produce short-term forecasts that are computed with the EM algorithm implemented with the first few eigenvectors ordered by descending eigenvalues. A cross-sectional bootstrap experiment is used to shed light on the sensitivity of the factor model to factor selection and to sampling uncertainty. The empirical number of factors seems more appropriately set through an analysis of eigenvalues, bootstrapped eigenvalues or the BIC than with more sophisticated information criteria. Confidence intervals derived from bootstrapped forecasts show the extent to which the data composition can support the hypothesis of business cycle co-movements and the selected factors can account for those shocks. Pseudo real-time out-of-sample forecast experiments conducted with a dataset of about two thousand series covering the euro area business cycle show that the SLID factor model outperforms benchmark models (AR models, leading indicators equations) for one-, two- and three- quarters-ahead forecasts of GDP growth. The accuracy of coincident forecasts compared to final estimates is not significantly different from Eurostat Flash or first estimates and is slightly superior to that of CEPR Eurocoin.

Keywords: bootstrapping, approximate factor model, GDP forecast, principal component analysis, EM algorithm, common factors.

* **Acknowledgements**: the author thanks M. McCarthy, C. Gayer, C. Denis, R. Liska and W. Roeger for useful discussions, R. McKenzie (OECD) for providing vintage euro area GDP series and Y. Bouquiaux for data management assistance. Shortcomings and errors are the responsibility of the author alone.

D. Grenouilleau is an economist at the European Commission, Directorate general for Economic and Financial Affairs (in the Unit *Econometric models and medium-term studies*). Please send comments to Daniel.Grenouilleau@cec.eu.int.

NON TECHNICAL SUMMARY

The modelling of business cycle developments and the short-term forecasting of GDP using leading indicators equations is usually not very robust to sampling uncertainty. First, leading indicators can only reflect specific shocks well. If shocks of a different nature occur in subsequent periods, then the indicators have to be reselected. Secondly, while numerous indicators are available, there is no rule that robustly prescribes which indicators should be selected. Moreover, results obtained using standard regressions are unfortunately very sensitive to the choice of predictors.

A factor model can provide a better response to these challenges. In the case of the approximate factor model described in this paper, there is *ex ante* very little selection from among all potential leading indicators. All series assumed to contain information about the current and/or future economic situation of the euro-area economy can be selected. The model distils from the pattern common to all leading economic series a signal about the business cycle in the near future that is cleansed of noise and idiosyncratic patterns of the series. Short-term forecasts about the business cycle can be directly derived from an efficiently extracted signal. In contrast to conventional forecast models based on regression on a few leading indicators, the forecast accuracy should be more robust over time, even if economic shocks of a different nature occur given that all potential indicators are used (at several lags).

The standard assumption of a factor model is that all indicators have two components: a common component corresponding to the general economic situation (business cycle) and an idiosyncratic component that is specific to each indicator. Following the methodology of Stock and Watson (1998), it is possible with principal component analysis (PCA) to extract from a large set of data common factors that summarise the unobserved common component to all series. Given that factor extraction is performed on a variety of series from countries of the euro area, the common component to all series reflects the overall business cycle of the euro area and can provide a good proxy for euro area GDP. Consistent estimates of the "true" (latent) common factors driving the business cycle can be obtained with large numbers of indicators and observations. In order to meet those requirements, about two thousand time series, covering various data from the twelve countries of the euro area, were collected.

The appropriate number of factors entering the model remains to a large extent an issue to be resolved by empirical applications. In a sense, the problem of indicator selection is replaced by that of factor selection, but economic judgment is of little help in this context, since factors are difficult to interpret directly. Moreover, the properties of PCA estimators in a non-asymptotic context are not known. This holds, in particular, for the effect of sampling uncertainty on the estimated factors.

Since PCA is a non-parametric method, bootstrapping provides a useful tool to shed light on these issues. The bootstrap is a resampling technique that is used to measure the sensitivity of model estimates to sampling uncertainty. The implementation of the bootstrap in the cross-sectional dimension of the data is particularly well suited to the issues raised by the consistency of approximate factor models. A monitoring of the variability of model estimates with bootstrapped samples (sample replicates of the same size randomly drawn with replacement from the original dataset) shows how consistent the factor estimation is. The intuition behind the use of a cross-sectional bootstrap is simple. Where principle components are consistently estimated in a factor model framework, they account for some common shocks driving the data. Then, resampling the data should only cause minor variations in factor estimates, given that any sample replicate should by definition exhibit the same common shocks. Conversely, large variability of factors (and forecasts) is likely to signal that the factors were not consistently estimated from the data with PCA or that the number of factors was not correctly set.

With cross-sectional bootstrapping, we obtain confidence intervals for the forecasts, which show the extent to which the data composition can support the hypothesis of business cycle co-movements and the selected factors can account for those shocks. Bootstrapping can also be used to select factors based on latent roots (eigenvalues corresponding to the principal components). The sensitivity of the forecast accuracy is checked for the other elements of calibration required by the model.

Empirical simulations, using about two thousand leading indicators and taking into account their real time availability, show that our approximate factor model outperforms benchmark models (simple stochastic model, leading-indicator equations) for one-, two- and three- quarters-ahead forecasts. The accuracy of coincident forecasts of GDP growth compared to final estimates is not significantly different from Eurostat Flash or first estimates and is slightly superior to that of CEPR Eurocoin.

GDP is one possible measure amongst others of business cycle conditions. This series is itself an imperfect estimate of business cycle conditions (due for example to seasonal and trading-day adjustment or measurement errors). All the discrepancy between measured GDP and the factor model coincident forecast should not be viewed as problematic since the former might not necessarily be a better reflection of business cycle conditions than the latter.

1. INTRODUCTION

This paper builds on previous research presented in Grenouilleau (2004), which aimed at the construction of a robust approximate factor model for short-term forecasting of business cycle series. The objective is to improve the robustness of the model, to monitor more precisely its forecast accuracy and, in particular to study in greater detail the issue of the selection of factors and the sensitivity of the model to the calibration.

◊ *General description of the approximate factor model*

The general framework of the model is factor analysis, the objective of which is to extract common or summary information from a large number of series. Following Reichlin (2002), each variable in the model is "represented as a sum of a component which is common to all the variables in the economy and an orthogonal idiosyncratic" component (residual). If most variables display co-movements, a few factors will account for a large share of the data variance. In the case of economic time series, the patterns of which usually reflect business cycle fluctuations, it can be expected that a few macroeconomic shocks reflected in the common factors will account for a substantial share of the data variance.

The approximate factor model used in this paper is analogous to the SLID factor model[1] introduced in Grenouilleau (2004): common factors, reflecting cyclical co-movements across series, are not estimated with a maximum likelihood method but with principal component analysis (PCA) as in Stock and Watson (1998). In theory[2], common factors are consistently estimated with PCA for N (predictors) and T (time observations) going to infinity. Principal components are extracted from a large set of series stacked at several lags. Forecasts are recursively computed with the EM algorithm, based on a simple projection of the variable of interest on the first few eigenvectors (sorted according to the descending order of eigenvalues).

◊ *The calibration*

Where the number of predictors exceeds the number of observations, the estimation of a factor model is not feasible with maximum likelihood[3], but only with PCA. With PCA, the use of a greater number of predictors can potentially lead to a gain in forecast accuracy provided that the constraint on their number is relaxed. On the other hand, parametric tests are no longer available in order to inform calibration choices, in particular the choice of the number of factors. In other words, more information is potentially available with PCA, but the robustness of the model may remain a source of concern, since factor estimation and forecast accuracy depend on calibration choices.

The use of approximate factor models with large N (greater than T) in empirical forecasting applications unavoidably involves a sizeable amount of calibration. The number of latent factors entering the model for a given variable of interest (here euro area GDP growth) remains to a large

---

[1] Sorted Leading Indicators Dynamic factor model, of which a special case is the Stacked Leading Indicators Dynamic factor model, which is dealt with in this paper.

[2] Cf. for instance Bai (2003).

[3] An efficient maximum likelihood estimation of a factor model (usually performed with a Kalman filter) requires a much larger number of observations than predictors in practice.

extent an unresolved issue in a non asymptotic framework[4]. Standard applications usually motivate the choice of the number of factors with adjusted in-sample fit[5], but the robustness of the calibration, for example, to the data composition or to the time sample, remains an open issue. Other "hyper-parameters" also need to be calibrated: the number of lags for the predictors (or the factors[6]), and the sample span (time window), which is used to estimate the model[7].

The reliability of the factor model is likely to be closely linked to the robustness of the calibration.

◊ *Model uncertainty*

In theory, approximate factor models allow us to be somewhat agnostic about the informational content of the many predictors available, insofar as all of them could be used (in case of doubt) without any risk of instability in the model estimation. Technically, it is possible to include as many predictors as desired, since there is no restriction on the number of series compared to the number of observations[8]. However, if the number of factors is not precisely known in practice, the uncertainty about the choice of factors is to some extent similar to that about the choice of predictors in small-dimension systems (OLS regressions or VAR). The issue of the selection of predictors is transferred to the stage of factor selection, even if circumvented at the stage of the data selection. The uncertainty at the stage of factor selection is all the greater given that decision rules based on the fit might involve similar methodological[9] or econometric[10] risks to those involved with fitting leading indicators to a given explained variable. Even worse, the choice of factors is more difficult than that of predictors in the case of low-dimension systems, since there is no direct mapping of factors to economic indicators and, hence, little economic interpretation available to inform the choice of factors as predictors.

Thus, the robustness of approximate factor models with a large cross-section dimension depends heavily on the decision rule for the selection of factors. The focus of this paper is to introduce a methodological framework based on a bootstrap experiment, which allows us to assess whether the calibration is robust to the uncertainty regarding the choice of the series entering the input database. The robustness of the model to data sampling is obviously not a sufficient condition to ensure that the calibration is valid for any time sample including future observations. However, the use of a cross-sectional bootstrap should provide an answer on two issues: is PCA estimation of factors consistent with a large cross-section and a fixed (small) number of observations? and is the calibration (and in particular the selection of factors) robust?

---

[4] Available asymptotic information criteria (e.g. Bai and Ng (2002)) do not seem applicable. See below the results of the implementation of such tests on empirical data.

[5] Information criteria are traditionally based on in-sample fit adjusted by a penalty term. Cf. Bai and Ng (2002), Stock and Watson (1998).

[6] In the case of the standard ("fitted") approximate factor model, cf. Stock and Watson (1998).

[7] It also concerns the filtering threshold, which calibrates the *ex ante* removal from the dataset of predictors with low correlation to the benchmark series in Grenouilleau (2004), but not in the model specification used in this paper.

[8] In contrast to parametric factor models estimated with maximum-likelihood, which require a non singular (T,N) data matrix.

[9] An *ad hoc* calibration to achieve a good out-of-sample fit is potentially subject to the same methodological problem as in-sample fitting. There is no guarantee that the fit obtained within a given sample can be replicated with additional observations. On ex ante vs. ex post forecast accuracy evaluation, see Clements and Hendry (2005).

[10] Overfitting, spurious regression, multicollinearity or other causes of non robust specification over time.

2. ECONOMETRIC FRAMEWORK

2.1. FORMAL PRESENTATION OF THE APPROXIMATE FACTOR MODEL

2.1.1. General presentation

The Stacked Leading Indicators Dynamic factor model is based on the Sorted Leading Indicators Dynamic factor model introduced in Grenouilleau (2004), which itself draws on the approximate factor model framework developed by Stock and Watson (1998). The aim of the models is to robustly extract the information contained in a large cross-section in order to forecast a variable of interest (namely euro area GDP) up to a few quarters ahead. The model is a "pure" factor model and not a "fitted" factor model[11]. GDP is included in the dataset of predictors and the projection of GDP on a subspace of factors is automatically derived from GDP loadings obtained through PCA. A very large number of predictors is used (about two thousand) and, moreover, all predictors are included at several lags in the dataset in order to extract *ex ante* the temporal dynamics[12] conveyed by the data. The EM algorithm is used to tune the GDP forecast, which is estimated just like a missing observation in the dataset[13]. Forecasts at subsequent steps ahead are recursively based on predictions obtained at previous steps. For the baseline model, no filtering of the input variables (predictors), such as that described in the original paper introducing the SLID model[14], is performed (the usefulness of filtering is examined in the section on analysis of robustness).

2.1.2. Factor extraction

Series are stacked at various lead[15] or lags into the data matrix:

$$X = [X_{-1}, X_0, X_{+1}, X_{+2}, X_{+3}, X_{+4}]^{16}, \qquad (1)$$

[11] In the standard approximate factor model, e.g. as introduced by Stock and Watson (1998), forecasts are derived from the estimation of an appended system including a subset of factors and the series to be forecasted.

[12] In contrast to standard approximate factor models, such as Stock and Watson (2002b), in which principal component analysis is performed statically (on coincident data only) and factors are bridged with lags to the variable of interest.

[13] See the full description of the algorithm in annex 1.

[14] The original model formulation in Grenouilleau (2004) allowed the dataset to be trimmed according to the cross-correlation of the predictors with the variable of interest (i.e. GDP). Predictors at a given lag with a correlation to GDP lower than a preset threshold were removed from the dataset at the same lag. Here, series are stacked but not sorted according to their correlations to GDP or, equivalently, the correlation threshold is set to zero.

[15] Series introduced with a lead in the data base are rarely used for GDP forecasting due to a lack of timely availability of such series. However, it is possible for a "coincident" forecast of GDP performed at the same date as the Eurostat Flash estimate release to use the first survey observations available for the subsequent quarter.

[16] All series are potentially assumed to contain leading information at a horizon of i quarters. They are thus shifted by i quarters into the future in the matrix X_{+i}. Note that the splitting of the matrix suits the objective of forecasts performed up to a theoretical maximum horizon of h=4 quarters. Further stacking might improve the model forecast performance at more remote horizons for some variables of interest, for which long-leading indicators are available.

where X_0 is a matrix of Q-Q_0+1 observations (rows) from quarter Q_0 to quarter Q and N series (columns), X_{+i} is the matrix of the same N series in columns but the observations are shifted by i quarters (downwards) into the future. In real time, X contains some missing observations at certain leads or lags due to the timeliness of availability of the series.

Let us consider a forecast for the horizon h (by convention h=0 for a coincident[17] forecast). The data matrix X is trimmed in the time dimension in order to retain X^h a time window of T quarters from quarter Q+h-T+1 to quarter Q+h, the last observation of which corresponds to most remote quarter Q+h to be forecasted. All the series that display a missing observation (at a given lead or lag) over this time-span are removed from the set X^h (at the relevant lead or lag) leaving N_h variables (columns).

The data are assumed to be generated by a k-factor structure:

$X^h = F_k^h.\Lambda_k^h + U^h$, where F_k^h contains k latent factors (columns) (2)

F_k^h and Λ_k^h are estimated with principal component analysis by solving the following recursive equation:

For i from 1 to T+h-1, given $F_{i-1}^h = \left\{ f_j^h \right\}_{1 \le j \le i-1}$ and $\Lambda_{i-1}^h = \left\{ \lambda_j^h \right\}_{1 \le j \le i-1}$ where i>0:

$$Min_{f_i^h, \lambda_i^h} \left\{ (X^h - F_i^h.\Lambda^h)(X^h - F_i^h.\Lambda^h)' \right\}, \tag{3a}$$

the solution[18] to which is to set F^h (T+h rows and columns[19]) to be the T+h eigenvectors of the variance-covariance matrix $\hat{\Sigma}_{XX} = \frac{1}{T+h} X^h.(X^h)'$ (3b)

The corresponding loading matrix Λ^h (N_h rows, T+h columns) is derived from:

$$\hat{\Lambda}^h = (\hat{F}^h)^{-1}.X^h \tag{3c}$$

The k latent factors of the model entering equation (2) are assumed to be the first k eigenvectors of $\hat{F}^h$ (by order of descending eigenvalues). Accordingly, only the first k rows $\hat{\Lambda}^h$ actually enter the latter equation (2).

2.1.3. Forecast computation

Let us assume that our variable of interest, euro area GDP, is included (coincidently and possibly with lags) in the data matrix. For convenience, let the coincident series of GDP be denoted by the last series in X^h. According to our definition, the last observation of GDP (and only the last) is missing at the coincident forecast horizon. The omission is the same several steps ahead irrespective

[17] "Coincident", one-quarter and two-quarters-ahead forecasts refer, respectively, to forecasts produced less than 3 months before the release of Eurostat GDP flash estimate, between three and six months ahead of Eurostat GDP flash estimate release and between six and nine months ahead. Let us recall that Eurostat's GDP flash estimate is released about 45 days after the end of the quarter estimated, hence the use of the term coincident for information available shortly before its release.

[18] See for instance Stock and Watson (2004).

[19] It is not necessary to increase the time period span by h observations at the horizon h compared to the coincident (h=0) forecast. In the following empirical applications, the time window span is in fact kept constant (constant T+h), meaning that T decreases by one observation for one more step ahead in the forecast horizon. T+h = 30 quarters in most numerical applications presented in this paper. The time window is thus shifted by h quarters for h-step-ahead forecasts compared to the coincident forecast.

of the forecast horizon, since forecasts at horizons $h_i>0$ can be recursively based on the previous horizons' forecasts h_j for all j<i.

$$\mathrm{GDP}_{\mathrm{T+h}} = X^{h}_{N_h,T+h} = \sum_{i=1}^{k} f^{h}_{T+h,i}.\Lambda^{h}_{i,T+h} \tag{4}$$

…is the forecast to be computed with k factors, *i.e.* the one and only missing observation in the data matrix X^h. The EM algorithm allows us to obtain joint estimates of f^h and Λ^h, that will rapidly converge to a unique solution, whatever first guess is used for $\mathrm{GDP}_{\mathrm{T+h}}$ in order to fill the missing observation and make the extraction of principle components possible over the whole time-span of T+h observations (for more detail about the implementation of the EM algorithm, see annex 1).

The EM algorithm requires an exogenous assumption regarding the number k of principal components, assumed to be estimators of the latent factors (see equation 4), which can be used to forecast GDP.

◊ *Motivation for such a specification*

Standard approximate factor models[20] generally involve the additional estimation of an appended OLS or VAR system including a selection of factors and the variables of interest. This can be avoided with our specification. The problem with appended systems is that an additional source of forecast volatility is introduced through the estimated linkage between the factors and the dependent variable. One should consider that factors are already estimated with error through PCA. The OLS or VAR estimation of the relation between GDP and the factors (potentially including lagged terms of both exogenous and endogenous variables) necessarily adds more uncertainty to the estimation of the coefficients (parameters), which can only be reduced through the use of a longer time span. Where samples with many observations are required, fewer series are available. Moreover, the likeliness of structural breaks in the input series or endogenous series is increases. Last but not least, more principal components are extracted from the data[21], and the number of potential combinations of factors rises exponentially with the number of principal components available. All in all, longer time spans do not necessarily increase the robustness of the linkage between factors and the modelled variable, nor do they enhance the informational content of the input data.

The SLID specification offers another trade-off between all these constraints: more series are available since the time window is narrower, fewer principal components are extracted, and the calibration essentially concerns the number of factors, not the additional estimation of the linkage between factors and the forecasted variable.

2.2. THE BOOTSTRAP METHOD APPLIED TO FACTOR MODELS

2.2.1. The bootstrap method

Bootstrapping is a computer-based method for assessing the accuracy of any statistics derived from a data sample. It is particularly useful in the case of non-parametric models, where confidence

[20] E.g. Stock and Watson (1998).

[21] The number of principal components is equal to the smallest dimension of the dataset, which normally corresponds to the number of observations in approximate factor models and to the number of series in maximum-likelihood factor models.

intervals for the estimates cannot be derived from the model. It is based on resampling the data randomly with replacement in order to construct approximated confidence interval for a statistic, which is function of random samples with an unknown probability distribution[22].

Suppose a random sample X of size N is observed from an unknown probability distribution F:

$F \rightarrow (x_1, x_2, ..., x_N) = X$, and θ a parameter of F: $\theta = t(F)$.

Bootstrapping uses the empirical distribution function $\hat{F} \rightarrow (x_1*, x_2*, ..., x_N*)$, which attaches probability 1/N that each x_i will proxy θ with $\hat{\theta} = t(\hat{F})$ according to the "plug-in principle".

A bootstrap estimate of the bias of $\hat{\theta}$ is:

$$\text{Bias}_F(\hat{\theta}) \approx Bias_{\hat{F}}(\hat{\theta}) = \hat{\theta}*(.) - t(\hat{F}) = \sum_{b=1}^{B} \frac{1}{B}\left[\hat{\theta}*(b)\right] - t(\hat{F}),$$ [23]

where $\hat{\theta}*(b) = s(x*^b) = s\left[(x_1*, x_2*, ..., x_N*)\right]$ is a bootstrap replication of $\hat{\theta}$ based on a bootstrap sample of the same size consisting of N data values drawn with replacement from the observed sample X. This Monte Carlo estimate converges to the expectation of s(x*) according to the law of large numbers.

A good approximation of the standard error of $\hat{\theta}$ can be obtained with the following bootstrap estimator:

$$SE_F(\hat{\theta}) \approx SE_{\hat{F}}(\hat{\theta}) = \sqrt{\sum_{b=1}^{B} \frac{1}{B-1}\left[\hat{\theta}*(b) - \hat{\theta}*(.)\right]^2}.$$

Under standard circumstances, the distribution of $\hat{\theta}$ becomes more and more normal as N grows large according to the central limit theorem, with a mean near θ and a variance near $\left[SE_{\hat{F}}(\hat{\theta})\right]^2$.

2.2.2. Monitoring factor consistency with cross-sectional bootstrap

◊ *The consistent estimation of factors with PCA*

Approximate factor models theoretically permit the consistent estimation of factors with a large number N of input series (predictors) based on principal component analysis (PCA). However, very little is known about the convergence of principal components to common factors in empirical applications with existing (non-simulated) data, i.e. without strong assumptions about the underlying data generating process (DGP)[24]. In an empirical setup or in a semi-asymptotical framework[25], a certain degree of uncertainty affects the estimation of factors with PCA: factors are approximated by principal components. Moreover, there is no definitive criterion for identifying the exact number of latent factors characterising a given set of data.

[22] For a general presentation, see Efron *et al.* (1993).

[23] The convergence is faster with a better bootstrap estimate, which replaces the second term (parameter estimate with the observed sample) with the average of bootstrap resampling vectors, see p. 130-133 in Efron et al. (1993).

[24] Cf. Bai and Ng (2002).

[25] For instance: with very large N and finite (small) T.

This theoretical issue may have important empirical implications: the composition of the (input) data might affect the factor estimation and forecast estimation[26] in a semi-asymptotic framework, and adding more data might not necessarily yield factors, which carry a signal that is more useful for forecasting any endogenous series[27]. If factor estimates or forecast estimates were sensitive to the choice of input series, this would challenge either the model specification or the econometric technique used. In principle, factors contain only a signal common to all the input business cycle series. If this common signal varies substantially across samples, it could mean that the sample replicates contain predictors that are irrelevant for describing the business cycle (case of mis-specification by analogy to OLS). Alternatively, it could mean that consistency requirements for factor estimation with PCA are not met with too small N and/or T.

PCA is a statistical method which is non parametric. As such, it gives no hint about the properties of the estimates or about confidence intervals around these estimates in a non-asymptotic framework. Only bootstrapping can provide such estimates and help to detect a problem of consistency in the estimation.

◊ *Factor consistency and cross-sectional bootstrapping*

In the case of a time series model estimated with PCA, the bootstrap method can be applied to the cross-sectional dimension of the data, considering that the dataset used is just one sample drawn randomly from a larger population of potential predictors. Bootstrapping allows us to obtain empirical distributions for any model estimate (eigenvalues, eigenvectors, loadings and forecasts[28]), which should converge in most circumstances to their true distributions under uncertainty regarding the choice of input series. The use of bootstrapping to analyse the stability of principal component analysis estimates is not new[29], but its application to time series models is a somewhat more recent and controversial issue. In a cross-sectional framework, the time sample used is non-random; only the cross-sectional composition is random and the problems of bootstrap estimation, which arise for time series models[30], are not relevant for this framework.

The intuition behind this bootstrap experiment is simple: by resampling the data, the bootstrap provides an assessment of the instability of the estimates arising from the selection of specific input series in a non-asymptotic framework or in empirical applications. According to the canonical approximate factor model framework, common factors are consistently estimated with large cross-sections[31]. A bootstrap helps us to assess whether factor estimates provided by principal component analysis are effectively consistent with N, given that common factors summarise the signal conveyed by all series, which is by definition insensitive to the inclusion of specific business cycle indicators in each data resample. In other words, one can check that the dataset is sufficiently large to disentangle common signals from idiosyncratic signals. If this consistency property remains valid in a semi-asymptotic framework (fixed T but very large N), the common factors should be identical across bootstrap replicates. But the other principal components (by order of descending eigenvalues) are likely to exhibit more variability across replicates, given that bootstrap replicates will not contain the same subsets of series capturing idiosyncratic shocks. Eigenvalues and loadings

[26] Cf. Grenouilleau (2004).

[27] Cf. Boivin and Ng (2003).

[28] In the baseline factor model presented in this paper, forecasts are obtained through direct projection of the benchmark series onto the latent factors.

[29] Cf. Beran *et al*. (1985), Daudin *et al*. (1988).

[30] Cf. Härdle et al (2003).

[31] The literature has not yet tackled the issue of the consistency requirements for N given T in a non-asymptotic framework. The issue of the time dimension of the dataset is neglected in the following developments.

(which give a unique projection of each series on the sub-space of principal components) should also be stable in the former case and more volatile in the latter. Forecasts that are computed as the sum of the factor observations (eigenvectors) weighted by the respective loadings should also be stable where consistent principal components are used.

In the approximate factor model framework, only common factors (consistently estimated principal components) should be used to forecast the benchmark series, while the use of idiosyncratic patterns for the modelling of an endogenous series is left to standard stochastic models. Low standard errors (SE) of bootstrap estimates are likely to signal that only latent factors were used in the estimation, and that they have been consistently estimated with principal components (the cross-section is sufficiently large and the dataset is consistently built to make the estimation converge). On the other hand, larger SE of the estimates suggest that either the cross-section is not large enough to obtain consistency or that the principal components used in the model do not correspond to common factors.

2.2.3. Monitoring forecast uncertainty with cross-sectional bootstrapping

An indirect application of cross-sectional bootstrapping provides a more accurate analysis of the robustness of the model estimation, and in particular model forecast accuracy, to the calibration. For each calibration set, the bootstrap produces estimates of the model statistics (e.g. out-of-sample forecasts and RMSE) and their standard errors. These estimates are unbiased irrespective of the composition of the input data and their SEs reflect the behaviour of the model with various calibrations carried out under uncertainty regarding the selection of predictors. This type of "model uncertainty" relates to the fact that we have access in real time to thousands of economic indicators but we cannot know for sure which indicators are relevant predictors[32] in the model. Monitoring the model forecast performance under a variety of calibration settings with bootstrapped statistics is likely to be more reliable insofar as the results are completely insensitive to data mining.

However, it should be stressed that other sources of forecast uncertainty - not just the uncertainty related to the selection of input series but that arising, for instance, from the selection of a specific set of observations[33] used to perform this exercise - are not taken into account in the cross-sectional bootstrap framework. This paper does not explore time-period bootstrap; such a task (left for future research) is not unfeasible but requires a complex randomisation of time observations given the dynamic structure of the model[34]. This shortcoming is compensated by the use of a long out-of-sample time span (25 quarters) and by the systematic monitoring of bootstrapped results on a quarter by quarter basis or with sub-samples. This is done in order to allow the detection of any problem of robustness of the calibration to the time sample used.

---

[32] With leading indicators equations, model uncertainty relates to the fact that the selected predictors are not necessarily the "true" ones, either because our economic judgement is wrong or because the decision rule for selecting regressors does not exclude the possibility of data snooping.

[33] Cf. parameter uncertainty in regression models. Stock and Watson (2004) mention temporal instability in their discussion about factor model instability and the bootstrap method could also be used to tackle this issue. Conversely, a bootstrap with random time observations and fixed data composition would provide estimates that are not necessarily robust to data snooping problems, especially with large but not very large cross-sections.

[34] In practice, time-period bootstrapping could be used to monitor forecast confidence intervals. However, monitoring the stability of eigenvalues or eigenvectors with such bootstrap framework is a complex task. One of the advantages of factor models is that factors and loadings are allowed to change over time, thus potentially adjusting to structural changes. This feature is highlighted through the use of a narrow time span (7½ years for the model presented here). Resampling the data in the time dimension might only result in a deterioration of the robustness of the estimates to structural changes, due to mixing up observations from different time periods (e.g. with moving blocks bootstrapping).

Bootstrap distributions indirectly show which calibration settings lead to lower RMSE and narrower confidence intervals based on bootstrap standard errors. Conversely, it is possible to detect calibrations under which the model accuracy is more sensitive to model uncertainty.

3. EMPIRICAL RESULTS WITH THE BASELINE CALIBRATION

3.1. THE DATA

3.1.1. Data processing

The various steps followed in the data processing are identical to those in Grenouilleau (2004)[35]. A very large set of data covering the twelve countries of the euro area and the euro-area aggregates is utilised. The series' selection criterion is judgmental: the selected series are assumed to contain information about the current or future economic situation and, therefore, to be possible leading indicators for the euro-area economy. An emphasis is put on good quality data in terms of accuracy and information content concerning the business cycle. A selection is made where many series are available for the same indicator.

All series undergo a homogeneous transformation: they are seasonally adjusted with Tramo Seats (except where only adjusted data[36] are available), converted to quarterly frequency, and transformed into first differences (a simple difference is used if the series can be negative, otherwise the difference in logarithms is used) in order to be made stationary. The real-time use of the model required a special treatment of monthly data. Where only one or two monthly values are available in a given quarter, a substantial loss of information would ensue if the latest observations of the series could not be used. Quarterly averages are thus replaced by rolling three-month moving-averages[37]. Series are stacked at several leads and lags (maximum three lags). Indicators for which no data are available for the quarter needed to compute forecasts are dropped[38].

The forecast computation is then derived from an optimisation scheme based on the first few principal components of the selected series (see the model presentation in previous section or the EM algorithm description in annex 1).

3.1.2. The pseudo real-time out-of-sample design

The data used in the model are generally available from a few days to about four months after the end of the relevant month. With a database of about 2000 series, conducting the out-of-sample experiment for every month with the data effectively available at that time and adjusted using

[35] Putting aside the fact that no filtering of the input data is performed. The data processing is completely automatised in order to allow for high-frequency (monthly) updates of the model forecasts.

[36] This refers in particular to the data from DataStream. In others words, whenever non-adjusted series are available, these series are preferred and are seasonally adjusted using Tramo-Seats.

[37] In other words, it is as if the quarterly average series is lagged by one or two months in time. Since factor analysis is performed on the whole series displaced in time, only the relevant (common) coincident information is extracted.

[38] For example, a coincident forecast for 2003Q2 is computed with 2003Q2 values of series introduced coincidently, 2003Q1 values of series introduced with one quarter lag and 2002Q4 values of series introduced with two quarters lag, etc. Series introduced with a lead are dropped except during the 15 days preceding the release of Eurostat Flash Estimate, as some series are then available for the month following the coincident quarter.

historical seasonal coefficients would be a heavy procedure. Hence, the procedure of the out-of-sample exercise had to be simplified.

Vintage (as published at that time) values of the reference series, GDP, were used for each quarter of the out-of-sample. For the predictor series, data revisions are neglected (the effect is most likely negligible given the size of the cross-section). However, the date availability of the last observation is taken into account.

At the reference date, all the non-lagged series, which do not have an observation for the quarter to be forecasted, are removed[39] from the data subset. For lagged[40] series, a similar selection is performed. Where the lagged series do not have an observation available at the reference date (after lagging) for the quarter to be forecasted (h-quarters-ahead), they are also removed from the dataset used for h-quarters-ahead forecasts in all out-of-sample simulations.

Dealing with data availability in the case of monthly series transformed into quarterly series is slightly more complicated. As previously stated, it would not be optimal to exclude a monthly series until such time as all three months become available in a given quarter. For all quarters of the out-of-sample, we use the three-month moving average ending in the month of the latest observation. In other words, it is as if the series was lagged by one (or two) month(s) when two (or one) monthly values are still missing in a given quarter (respectively). As all series are introduced at several lags, the latter series are also introduced with 4 (or 5), 7 (or 8) month lags (respectively), etc.

The configuration just described corresponds to the information available in a real-time situation for any quarter to be forecasted in the out-of-sample span. Because of seasonal adjustment revisions, real time forecasts may not be exactly identical, but are most likely extremely close, since the number of series is very large.

3.1.3. Data sources

◊ *The benchmark series: euro area GDP*

Thanks to the OECD[41], vintage series for euro-area GDP dating back to 1999 have recently become available. It is therefore possible to construct an out-of-sample experiment starting in 1999 based on "vintage" series (meaning here that the series are identical to those published historically). The vintage series are not corrected for changes in the quarterly national accounts methodology that have occurred since 1999. As a result, some large revisions can be monitored, in particular for 1999Q1. Given the magnitude of the GDP growth revision (about 0.4 pp) for the latter quarter, the observation is not included in the out-of-sample span.

◊ *The predictors*

The dataset used to perform empirical experiments has changed compared to our previous paper[42] on the same model. Series from the monthly and quarterly Business and Consumer Surveys of the European Commission (2003) and from Comext trade data (in value terms) are the same. However, series covering industry and retail output data, housing starts, car sales, nominal data including

---

[39] More precisely, all the quarterly series with no value available for the quarter to be forecasted, all the monthly series with no monthly values available for the quarter to be forecasted.

[40] The method is the same for series introduced with a lead of one quarter.

[41] Vintage series published before 2002Q4 are unfortunately unavailable from Eurostat. In our previous paper (2004), we had recourse to personal GDP files collected from Eurostat at the time of release, dating back to 2001Q3.

[42] For a more detailed description, cf. Grenouilleau (2004). We focus here on differences.

industrial and retail prices, monetary aggregates, stock markets indices, nominal exchange rates, interest rates, labour market data, business start-ups, bankruptcies, and additional specific surveys are taken from Ecowin instead of the Datastream series used previously. Let us recall that the data are broken down to national and/or sectoral series where available. The composition of the data by country and their economic nature is summarised in the table below.

type	AT	BE	DE	EA	EL	ES	FI	FR	IE	IT	LU	NL	PT	_other	Total
Survey	42	42	71	80	42	50	42	143	42	59	9	42	42		706
Financial	5	6	37	37	3	10	5	8	8	24		13	6	3	165
Demographics		1	3	1		14	1	1		1		2			24
Labour	2	21	18	30	1	21	1	13		8		1	5		121
Output	6	11	13	5	10	17	12	20	7	28			12		141
Prices	10	22	54	20		53		37	10	42		44	8		300
Sales	1	2	6	4		5	2	1	2	22		11			56
Foreign trade	36	38	37		36	39	36	38	36	37		37	36		406
Total	102	143	239	177	92	209	99	261	105	221	9	150	109	3	1919

3.2. OUT-OF-SAMPLE RESULTS

3.2.1. Comparison with various benchmark models[43]

◊ *Comparison with other forecast models*

The following table summarises 25 quarters of out-of-sample RMSE for a selection of standard short-term forecast models available for euro-area GDP. Most models (except purely stochastic models) seem to deliver accurate forecasts in the second month following the end of the predicted quarter. This is notably the case with the OECD leading indicator equations[44] that include hard data. However, leading indicators are typically subject to data mining problems, which are for example revealed when series are revised[45].

Just a few months before the release of flash or first estimates, the forecast accuracy of leading indicators equations deteriorates rapidly and they no longer offer a reliable alternative to a parsimonious AR(1) model. On the other hand, the performance of factor models (SLID model and Eurocoin) is more stable. The accuracy of Eurocoin deteriorates significantly for one-quarter-ahead forecasts (RMSE at 0.28%). However, it remains relatively low two months ahead (RMSE at 0.25%), although the model was not designed to produce estimates more than a few weeks ahead of Eurostat Flash estimate[46].

---

[43] Detailed tables are displayed in annex.

[44] Cf. Sédillot and Pain (2003).

[45] An update of the statistics presented in Sédillot and Pain (2003) based on revised series would provide a more reliable assessment.

[46] Two or three months ahead forecasts are based on proxies: the rolling MA3 (three-month moving-average) of the Eurocoin index ending in the last month available is used as a proxy for the growth rate of the relevant quarter. All three months of the quarter are usually available two weeks before the release of the Flash estimate for the latter quarter. Note that the accuracy is slightly overestimated due to the fact that the statistics is computed with the final (revised) index and not with real-time estimates of the index subject to (slight) revisions.

Forecast RMSE (a) (in %)	SLID factor model	Eurocoin (b)	AR model	ECFIN former GDP indicator	OECD indicator equations (c)
Coincident	0.19 to 0.21	0.22 to 0.28	0.28	0.19 to 0.35	0.21 to 0.30
One-quarter ahead	0.21 to 0.24		0.34	> 0.35	0.31 to 0.39
Two-quarter ahead	0.24 to 0.31		0.36		
Out-of-sample span (d)	25 quarters 1999Q3-2005Q3	25 quarters 1999Q3-2005Q3	25 quarters 1999Q3-2005Q3	12 quarters 1998Q1-2000Q4	20 quarters 1998Q1-2002Q4

NOTES:

(a) The RMSE varies across months in which quarterly forecasts are produced, hence the ranges

(b) Based on 3-month moving averages of the monthly indicator

(c) Combination of hard/survey indicators based on 0 to 2 months of current quarter information

(d) Different spans are a source of bias for a direct comparison

The SLID factor model accuracy is slightly better than that of Eurocoin according to the Diebold-Mariano (DM)[47] statistics; the hypothesis of better forecast accuracy of the SLID model is associated with a confidence level of only 77% at the date of release of Eurostat Flash estimate, 90% for two-month-ahead forecasts and 97% for three-month-ahead forecasts.

At more remote horizons, the stability of the SLID factor model suggests that it performs an efficient extraction of the leading signal contained in economic series up to 6 months ahead of Eurostat Flash estimate release. The accuracy of the SLID model within this range of forecast horizons is rarely shared by other short-term forecast models. For three-quarter-ahead forecasts, the DM statistics rather allow us to reject with only 80% confidence the hypothesis of equal forecast accuracy with an AR model.

◊ *Comparison with GDP first estimates (or Flash estimates where available)*

A common feature of most models for coincident growth forecasting is that the RMSE over relatively long out-of-sample spans does not seem to get lower than 0.20% (for quarter-on-quarter growth). The main reason for this stylised statistical fact could be that real-time estimates of GDP are themselves not very accurate. The RMSE of first (or Flash where available) estimates compared to the latest estimates available is 0.20% from 1999Q1 to 2004Q2[48]. There is indeed no difference in accuracy between the SLID model and the Flash/first estimate with respect to forecasting the final estimate of GDP[49]. Changes in the methodology (seasonal, trading-day adjustments, index chain-linking) naturally play a certain role in those revisions but do not account for the bulk of the change[50].

One problem is that the GDP series (of which the last observations will be substantially revised) is used by short-term forecast models and the noise in real time is likely to give rise to a deterioration in the forecasts, specifications and calibrations of such models. Indeed, the SLID model forecasts

[47] See the description of the statistics and detailed results in annex 3.

[48] This compares to 0.21% for the coincident factor model forecast. These are final or "semi-final" estimates, given that the figures for the last quarters of the out-of-sample will be further revised. The revision process usually lasts about 8 to 10 quarters, putting aside changes in the methodology. The out-of-sample span used for this statistics excludes the last 5 quarters for which the latest estimate is obviously a bad proxy for the final estimate.

[49] The DM statistic is 58% over 1999Q1-2004Q2 and even 67% over 1999Q1-2005Q3. The accuracy of first/flash estimates forecasts of the coincident factor model and of the Eurostat latest estimate can also be compared; the DM statistics show again that there is no difference in accuracy.

[50] For example, the first estimate for 1999Q1 of 0.43% was gradually revised to 0.96% 9 quarters later and subsequently revised downwards to 0.85% with a negligible impact from changes due to index chain-linking or trading-day adjustments.

with more accuracy the first estimate of GDP (RMSE at 0.16%) than the final estimate of GDP. The reason could be that the first estimate of GDP for a given quarter is itself affected by errors in the estimates of preceding quarters (also bound to be revised over time).

Last but not least, the factor model estimate gives some indication on the direction in which the Flash/first estimate might be revised. The direction of revision is correctly forecasted in 77% of the cases (over 1999Q1-2005Q3), where the factor model's estimate is significantly different from the Flash/first estimate (at least 0.05 pp difference).

All in all, the factor model coincident estimate should be seen as equally accurate in statistical terms as Flash/first estimates, but with different properties. GDP flash/first estimates are systematically biased (downwards) and more volatile (probably more noisy). The factor model estimate is on average unbiased but might underestimate growth in periods of high growth and overestimate it in periods of low growth. Since the latter nowcasts/forecasts are also smoother, they also provide a signal about the position in the business cycle, which is easier to interpret insofar as it is somewhat cleansed of very high frequency fluctuations, due for instance to trading days or seasonal patterns.

◊ *Forecasts revisions*

An important issue for a short-term forecast model is the consistency of the revisions when more data become available. With the SLID factor model, forecast revisions have a relatively low magnitude compared to what is typically obtained with a system of different leading indicators equations (one for every forecast horizon). The root mean squared difference at 0.09% is quite low between 6 months and 3 months ahead forecasts and becomes very low (0.05%) between 3 months ahead and coincident forecasts.

Forecast revisions	Hit ratio: correct revision direction*	Root mean squared differences
One-quarter-ahead to coincident forecast	91%	**0.05%**
Two-quarter-ahead to one-quarter-ahead forecast	91%	**0.09%**
Three-quarter-ahead to two-quarter-ahead forecast	85%	**0.14%**

* Differences equal or larger than 0.05%

Last but not least, forecast revisions seem to be well behaved: the direction of the revision compared to the final estimate (that will become available several quarters later) is correct in 90% of the cases (85% for two-quarter-ahead forecasts).

3.2.2. Forecast robustness under uncertainty

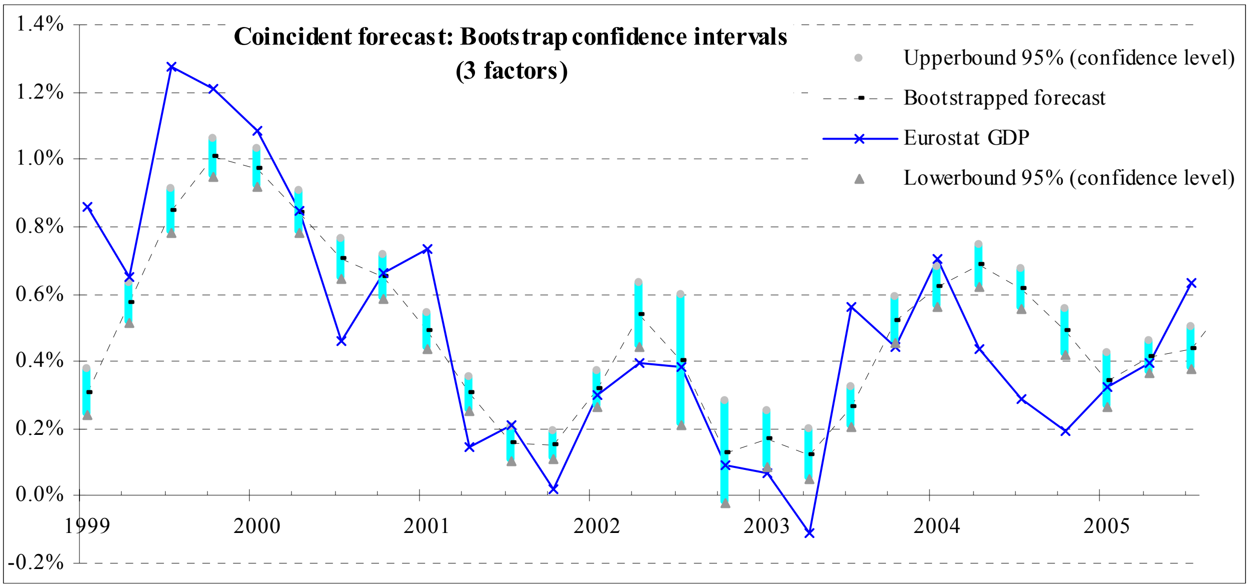
◊ *Confidence intervals for the forecasts*

An interesting application of cross-sectional bootstrapping to approximate factor models is to monitor how sensitive the forecasts are to the choice of input series. This experiment provides some insight about the convergence of principal components to factors, given that samples drawn randomly from the same database should by definition be associated with the same common factors. Large standard errors would signal either a too small number of series for the extraction to converge, the selection of inappropriate data or a wrong calibration of the number of factors chosen to perform forecasts.

Compared bootstrapped/sample forecasts accuracy				
99Q3-05Q3	Coincident	1Q ahead	2Q ahead	3Q ahead
1 factor	0.000%	0.000%	0.000%	0.000%
2 factors	-0.003%	-0.004%	-0.004%	-0.001%
3 factors	0.000%	0.000%	0.003%	0.000%
4 factors	-0.002%	0.000%	0.000%	-0.006%
5 factors	0.000%	-0.011%	-0.008%	-0.027%
6 factors	-0.010%	-0.024%	-0.014%	0.000%
7 factors	-0.016%	-0.019%	0.001%	-0.005%

A first important result is that bootstrapped forecasts are only very marginally different from the forecasts directly obtained with the original sample: there is no bias due, for instance, to data snooping. The use of bootstrapped forecasts improves the forecast RMSE very marginally only where more factors are used (negative difference in the table above between bootstrapped forecasts RMSE and sample RMSE). Under a full bootstrap experiment (including a time-period bootstrap), the property of unbiasedness would most likely still hold given the absence of bias for any quarter of the out-of-sample period.

The graph below[51] shows that the standard error (SE) of quarterly coincident forecasts is very low. For example, the SE is between 15 and 20 times smaller than average growth for coincident projections. The corresponding forecast range at a 95% confidence level is narrower than ±0.1 pp. However, the SE is much larger for very few specific quarters of the out-of-sample period, for instance 2002Q3 to 2002Q4. The explanation mostly lies in the fact that diverging patterns are displayed by the various subsets of data (e.g. survey data vs. output or industrial production data) and it is more difficult to extract a common pattern summarised by a few factors. A confirmation of the latter assumption can be found in the bootstrapped forecasts obtained with data excluding survey series (see next part for details).



A potential remedy to large bootstrapped SE can be, in some cases, to reduce the number of factors used in order to base forecasts on "more common" signals and reduce the bootstrapped SE. In other cases, the confidence range might simply reflect strong heterogeneity in the patterns of the various predictors and the only remedy would be to reconsider the choice of predictors.

◊ *Compared RMSE and confidence ranges with alternative choices of factors*

The results presented above correspond to the baseline calibration of the model, which includes a selection of three factors. Forecasts based on alternative choices of factors exhibit other features in terms of accuracy and dynamics. The RMSE corresponding to three or four factors are much better than those with one or two factors. With five, six or seven factors, they usually deteriorate steadily, but not for all horizons.

RMSE of the factor model bootstrapped forecasts				
99Q3-05Q3	Coincident	1Q ahead	2Q ahead	3Q ahead
1 factor	0.322%	0.342%	0.380%	0.408%
2 factors	0.308%	0.330%	0.387%	0.402%
3 factors	0.187%	0.207%	0.241%	0.308%
4 factors	0.194%	0.211%	0.244%	0.306%
5 factors	0.199%	0.223%	0.253%	0.316%
6 factors	0.215%	0.245%	0.262%	0.307%
7 factors	0.230%	0.250%	0.263%	0.310%

[51] See graphs corresponding to other forecast horizons and detailed results in annex 3.

The signal captured with less than three factors is clearly lagging GDP. With three factors or more, the index is highly correlated with GDP coincidently. This indicates that the third factor conveys coincident or leading information, whereas the first two provide mainly lagging information. Beyond 3 factors, the coincident correlation tends to decline marginally, while the index correlation with a lead of one quarter is improved; subsequent factors are likely to contain more leading (but also noisy) information.

Correlation of the coincident forecast index with GDP				
99Q1-05Q3	1Q lag	Coincident	1Q lead	2Q lead
1 factor	66%	33%	8%	-5%
2 factors	62%	35%	2%	-31%
3 factors	83%	80%	61%	27%
4 factors	82%	79%	66%	30%
5 factors	83%	78%	66%	32%
6 factors	81%	75%	65%	30%
7 factors	80%	72%	65%	27%

The graphs in annex 3 show that the more factors are used, the wider bootstrapped confidence intervals are. The explanation is simple: where additional factors are used, it is more likely that information pertaining to large subsets of (but not all) series is captured. The forecast variability according to the exact composition of the data is thus increased. Where few factors are used, there is a trade-off between the accuracy of the model, which requires using as many relevant factors as possible in order to retrieve more information, and the width of the confidence intervals, which is smaller if only common information is used. Very narrow confidence ranges are, for example, obtained with only one factor even for three-quarter-ahead forecasts. The problem is that the signal extracted with this factor is very flat, particularly at remote forecast horizons. More precise information needs to be found in the subsequent factors, the risk being that some factors reflect signals that are not common to all series. Beyond three or four factors, the RMSE no longer decreases, while the averaged SE increases rapidly. The use of additional factors raises forecast uncertainty and, although more information is available, the RMSE does not improve any longer.

SE of the factor model bootstrapped forecasts				
99Q3-05Q3	Coincident	1Q ahead	2Q ahead	3Q ahead
1 factor	0.017%	0.015%	0.013%	0.011%
2 factors	0.028%	0.031%	0.045%	0.101%
3 factors	0.038%	0.041%	0.054%	0.078%
4 factors	0.048%	0.053%	0.063%	0.085%
5 factors	0.055%	0.064%	0.076%	0.102%
6 factors	0.064%	0.075%	0.096%	0.103%
7 factors	0.076%	0.084%	0.111%	0.113%

These observations confirm the notion that factor selection is a key issue for the forecast accuracy of factor models. Cross-sectional bootstrapping also provides a view on how much the data support a hypothesis of common patterns for the number of common factors specified. The combined readings of RMSEs and forecast index correlations might also indicate that the relevant information about coincident GDP is not necessarily clustered in a simple manner in the first factors across out-of-sample quarters, especially for forecasts at more remote horizons.

4. AN ANALYSIS OF FORECAST ACCURACY[52]

The objective of a properly conducted out-of-sample experiment is to measure the sensitivity of models to forecast uncertainties. For example, a high out-of-sample RMSE combined with good in-sample fit suggests that the model is not robust to parameter uncertainty (large standard errors on parameter estimates), and/or to model uncertainty (incorrect model specification), and/or to future uncertainty (unsuitable explanatory variables for the future)[53]. The longer the out-of-sample time span and the less *ex post* information is used, the more reliable is the assessment of the model's robustness to forecast uncertainties. A fair assessment of calibrated models is more difficult where calibration settings are not completely independent from the *ex post* fit. It is thus useful to monitor the sensitivity of the model to calibration settings.

4.1. THE NUMBER OF FACTORS

Approximate factor models theoretically permit the consistent estimation of factors through PCA with a large number of predictors. However, little is known about convergence in a non-asymptotic framework, e.g. in empirical applications with non-simulated data. Principal components simply account for decreasing shares of the data variance (by descending order of eigenvalues) and, in practice, with a very large number of series, there is no simple criterion to determine which principal component is a factor and which is not. Calibrated information criteria recently provided in the literature do not seem robust to underlying assumptions on the data generating process. However, the selection of factors according to empirical PCA rules (based on eigenvalues) and supported with cross-sectional bootstrap findings, produces better out-of-sample results.

4.1.1. Bai and Ng (2002) information criteria (BNIC)

Bai and Ng (2002) formulate the problem of estimating the number of factors as that of model selection, each model allowing for a different number of latent factors. They introduce three information criteria based on the residuals of the time-series regressions of predictors on a given set of k factors corrected by a penalty term[54]. The suggested number of factors is always one for our empirical setup.

However, according to out-of-sample forecast results or eigenvalues distributions, the number of latent factors is certainly not one but rather three or four. Given that the BNIC shows satisfactory Monte-Carlo results with N and T of various magnitudes, it is likely that the penalty term is over-calibrated for non-simulated data. Business cycle series (especially where many different predictors are used) have a relatively low signal to noise ratio[55], at least much lower than that implied by the

---

[52] For a more simple discussion of the model accuracy across forecast horizons in this section, the coincident forecast is that performed at the same time as the release of Eurostat Flash estimate, the one-step ahead forecast three months later, and so on.

[53] Using the typology introduced in Garratt *et al.* (2003).

[54] See annex 4 for more details.

[55] The signal is even lower where series are lagged. However, the Bai and Ng IC, based only on coincident series, also suggests one factor.

DGP considered in the above-quoted paper. With noisy data, the penalty term does not seem appropriately scaled to the large residuals of the series' regressions on the factors, irrespective of the subset of factors used.

4.1.2. Bayes Information Criterion (BIC)

The BIC[56] based on the residuals of the GDP regression on the factors (instead of the residuals of the predictors' regressions) seems to be better behaved. Detailed results in annex 4 are broadly in line with PCA analysis of eigenvalues and out-of-sample results. The minimum score is generally obtained with three factors (sometimes two factors only) across out-of-sample quarters, and the average score (across the out-of-sample span) is minimised with three factors at all forecast horizons.

In this framework, the BIC outperforms the BNIC for a simple reason. We are looking for the optimal number of factors for the prediction of euro area GDP, assuming that all indicators gathered in our database contain some information about the variable of interest. The residuals of the regression on a subset of factors of some predictors that poorly correlated with GDP are necessarily less informative that the regression on a subset of factors of GDP itself. In other words, the accuracy of BNIC could be improved compared to the BIC only if all predictors contained a strong signal about GDP and very little noise, which is obviously not the case with a large number of noisy predictors.

4.1.3. The standard PCA approach: analysis of eigenvalues

The magnitude of the eigenvalues obtained with PCA is generally considered to provide important information in order select latent factors. The sum of the eigenvalues corresponding to a set of eigenvectors directly informs us about the share of the data variance which is accounted for by those eigenvectors (factors). In practice, particularly where no precise prior can be used (e.g. the predictors are very numerous or noisy), it is more customary to monitor the decay of the eigenvalues. For N series and T principal components, a solution would be to test the hypothesis that the T minus k smallest eigenvalues are equal. If the hypothesis is accepted, there is no point in allowing for more than k latent factors[57]. The problem is that the distribution of the statistics does not tend to a standard form with very large N. A standard practice in PCA is thus to monitor empirically the rate of decay in decreasing eigenvalues[58]. Eigenvalues close to one another are considered to signal uninformative principal components and, therefore, non latent factors.

Before examining the results quarter by quarter, an important remark should be made. Due to data availability constraints, the number of series available across the out-of-sample span is not the same. There are roughly twice as many series available (with 30 observations) for 2005 than for 1999, i.e. at the beginning of the out-of-sample simulation (many series were discontinued and are no longer available or have been replaced by shorter series). Thus, there is no guarantee that the number of latent factors is exactly the same over the whole out-of-sample time span (1999-2005).

There are two options for monitoring the decay of eigenvalues in our empirical setup. The first option is to monitor eigenvalues (also called "latent roots") quarter by quarter in the out-of-sample

---

[56] In Grenouilleau (2004), the results correspond to the BIC criterion (where the sum of squared errors is computed with a regression of GDP on the factors and not of the predictors on the factors).

[57] See Lawley et al. (1971), p. 20.

[58] The "scree test" is based on such methodology, cf. Cattell (1966).

period and the second option is to average across quarters the percentage of variance "explained" by each eigenvalue. Conclusions drawn from the second option framework are valid only if the latent roots (and factors) are structurally the same over the out-of-sample period, irrespective of the quarter and regardless of the number of predictors that are actually available.

◊ *Eigenvalues quarter by quarter in the out-of-sample*

The ratio of each eigenvalue to the sum of all eigenvalues is the percentage of variance accounted for by each factor. In order to monitor the decay of eigenvalues more easily, we focus on the difference between the percentage of variance explained by eigenvalue i and that for the next eigenvalue i+1. According to the Cattell (1966) criterion, latent factors extraction should stop before a latent root carries no significant gain in explained variance compared to the subsequent eigenvalue. The relevance of this empirical criterion can be easily checked with a Monte-Carlo experiment. Where the data are generated with a factor structure, the first latent roots corresponding to the factors (their number is equal to the number of factors) are much larger than the subsequent eigenvalues, which are broadly equal to each other. Conversely, where the data is generated completely randomly, all eigenvalues are broadly equal to each other.

On the graphs in annex 4, the first difference (for more clarity) of variance explained - stacked for all quarters of the out-of-sample - is displayed eigenvalue by eigenvalue. In our empirical application, the criterion usually prescribes that we keep the first three factors for coincident, one-quarter-ahead and two-quarters-ahead forecasts. The decision is perhaps more difficult in the case of three-quarters-ahead forecasts: in some cases only the first factor is obviously providing a gain in explained variance or, alternatively, one could choose the first four factors because of a slight drop in the percentage of variance explained after the fourth factor. More surprisingly, a further gap between the 26^{th} and 27^{th} eigenvalues appears for some quarters in the case of coincident forecasts, which might signal some explanatory power rooted in the 26^{th} eigenvector. This point is left for future research.

◊ *Averages of latent roots percentage across out-of-sample quarters*

Averages across quarters of percentage variance of eigenvalues can also be used in order to average out fluctuations due to the joint variations in the number of predictors and time-period idiosyncrasies. The decision based on the Cattell criterion is simplified: the gain in variance explained at all forecast horizons is on average much lower beyond the third factor and the criterion suggests keeping three factors. For three-step-ahead forecasts only, the decay in variance gain is perhaps slower and four factors could also be selected based on the criterion. The empirical performance of the model based on out-of-sample results is in fact only marginally worse with four factors at this forecast horizon.

◊ *Bootstrap analysis of latent roots*

The preceding results were obtained with a given, albeit very large, set of business cycle indicators. Under uncertainty regarding the choice of predictors, it might be interesting to check whether the monitoring of eigenvalues would suggest the same number of factors for all data resamples. Given that common factors should by definition reflect shocks that are common to all series, resampling the data should only cause minor variations in the properties of the latent roots. If this is not the case, either the eigenvector is not a common factor, or the factor estimation is not perfectly consistent insofar as variations in the composition of input data affect the extraction of the signal common to all series.

The "scree" test performed with bootstrapped eigenvalues' percentages averaged across quarters[59] shows again that the first three factors should selected; the fourth eigenvalue is usually not significantly different at a confidence level of 95%[60] from the fifth in the sense that both confidence intervals overlap. This result holds irrespective of the forecast horizon[61]. Common factors should be by definition the same across resamples. The bootstrap provides unbiased estimates of the eigenvalues: comparing ranges for the i^{th} bootstrapped eigenvalue and for the subsequent $(i+1)^{th}$ eigenvalue is nothing more than checking the gain in explained variance under uncertainty regarding data composition and sample estimation[62]. The advantage of the bootstrap methodology is to take into account the uncertainty associated with factor estimation. However, the derivation of a formal test is difficult given the correlation between the distributions of the eigenvalues.

The same exercise can be performed with bootstrapped eigenvalues quarter by quarter in the out-of-sample span. Similar conclusions can be drawn with respect to the number of latent factors for most quarters. However, some variations appear, in particular across horizons. For coincident forecasts, the gap between the 26^{th} and 27^{th} eigenvalues is supported for some quarters, which would signal some explanatory power conveyed by the 26^{th} eigenvector. On the other hand, bootstrapped eigenvalues cast doubt on the actual explanatory power of some factors for three-quarter-ahead forecasts; the gap between the second and third factors is no longer statistically significant for some quarters, nor is that between the third and fourth factors in some cases. A choice of three factors at this horizon could thus be challenged for some quarters. Large forecast confidence intervals are often associated with overlapping ranges for the eigenvalues. While a decision criterion based on averaged eigenvalues (across quarters) yields consistent empirical results, there is room for future research on the fine-tuning of factor selection.

4.2. THE DATA TIME SPAN

4.2.1. Potential issues raised by a narrow time span

The span of the time sample has been calibrated at 30 observations (7 ½ years), in order to match the duration of at least two euro-area short business cycles. The determination of the length of business cycles in the euro area has been complicated, especially in recent years, by the combination of volatile and lower growth. It appears from the graph opposite that GDP cycles have neither a constant duration (between 6 and perhaps 7 ½ years), nor a perfectly stationary pattern[63].

The use of a small time span does not intrinsically pose estimation problem[64]. A limited number of non-stationary predictors is unlikely to affect the estimation. However, estimation based on a

---

[59] More precisely, the ratio of each eigenvalue divided by the sum of eigenvalues is averaged across quarters of the out-of-sample period.

[60] Based on the empirical bootstrapped distribution.

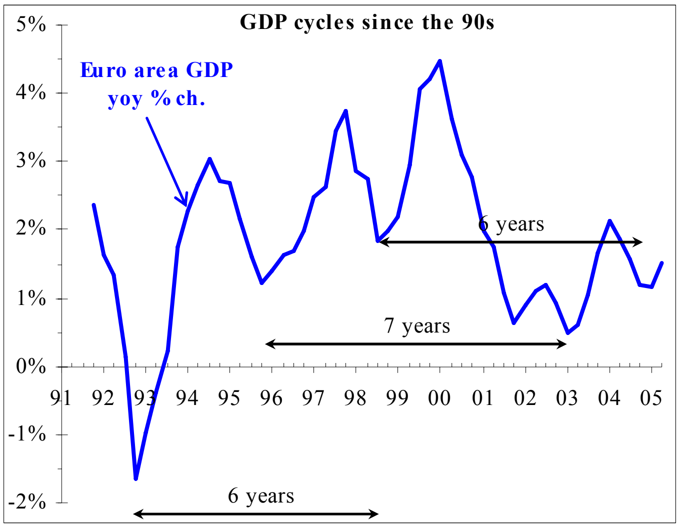
[61] See detailed results in annex 4.

[62] A more complete and rigorous framework for this bootstrapping exercise would be to randomise the time period choice. While the mean of bootstrapped eigenvalues across quarters is probably a good proxy, the SE is likely to be larger.

[63] The standard interpretations are that there is a break in the trend-stationary series, or that the series is not stationary; cf. Perron (1989).

[64] By contrast, the estimation of a bridge equation would not be very robust with a narrow time window, as the number of available observations would be too small compared to the number of indicators to yield estimated parameters with

significant share of non-stationary series might produce inconsistent results[65], as in the case of OLS, if the dependent variable is not cointegrated with the trended predictors.



Further investigation on this complex issue is beyond the scope of this paper, but it can be mentioned that samples having a smaller time span than the series cycles would most likely not be conducive to consistent estimation of factors and forecasts. The model implicitly assumes that the trending patterns of the series have some commonalities. A very short sample involves a higher risk of discrepancy between the sample trend in GDP growth and the average trend based on the series commonalities (which is necessarily smoother because some series display less pronounced cyclical patterns and because high-frequency idiosyncrasies are averaged out).

4.2.2. Empirical results with 24, 30 and 36 observations

Given that very little theoretical analysis is available, it is interesting to check empirically whether the use of a different time span affects the model accuracy. Experiments for coincident forecasts were conducted with spans of 24 and 36 observations. The number of eigenvalues is equal to the number of observations, where the cross-section dimension is larger than the time span. With a different number of eigenvectors, it is first necessary to check whether the same number of eigenvectors conveys the relevant information, i.e. business cycle signals. The analysis of bootstrapped eigenvalues suggests that the optimal number of factors over the whole out-of-sample period is the same for all time spans[66].

Forecasts results are similar with 24 and 30 observations[67]. The quadratic mean of the bootstrapped forecasts' SE is the same, meaning that confidence ranges have on average the same width. All in all, the increase in the number of series available[68] might compensate the loss of precision due to the use of a narrower sample span.

On the other hand, the model seems to be less accurate with a time span of 36 observations. The RMSE for 2000Q3-2005Q3[69] increases from 0.17% (24 obs.) or 0.18% (30 obs.) to 0.21% (36 obs.)

---

high confidence level. Moreover, there would be an increased risk that some indicators are not stationary over a narrow time window. The use of a panel estimation method reduces considerably the confidence intervals around forecast estimates due to uncertainty regarding the choice of predictors.

[65] The trended predictors are likely to dominate the first principal components.

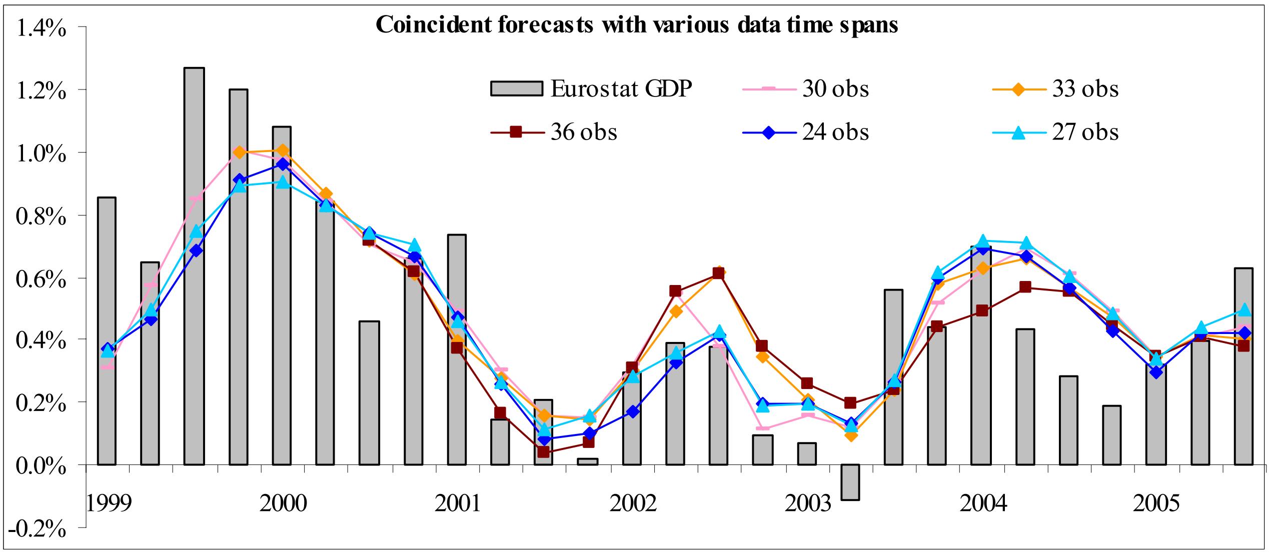
[66] Significant variations appear for some quarters of the out-of-sample period. The number of distinct eigenvalues is two rather than three for some quarters, with a 36 observations span.

[67] The coincident forecast RMSE is 0.21% with 24 observations vs. 0.19% with 30 observations. But the difference is mostly accounted by the first four quarters of the out-of-sample period. The respective coincident and one-quarter-ahead correlations are 77% and 56% vs. 80% and 61%, which indicates that the index of forecasts is slightly less correlated with GDP.

[68] Recall that about twice as many series are available at the end of the out-of-sample period with a time window of 7 ½ years, due to attrition in the first half of the 1990s. More series are thus available with a shorter time window.

[69] Out-of-sample results are not available for quarters prior to 2000Q3 with 9-year time windows due to the starting date of the database.

The reduction in the number of series available appears to outbalance the potential gain in accuracy linked to a more precise estimation of common patterns through a larger span of observations.



4.3. THE NUMBER OF LAGS FOR THE STACKED INPUT SERIES

RMSE of the factor model with various leads/lags of stacked series					
	Coincident	1Q ahead	2Q ahead	3Q ahead	4Q ahead
1 lead	0.270%	n.a.	n.a.	n.a.	n.a.
coin + lead	0.254%	n.a.	n.a.	n.a.	n.a.
coin	0.247%	0.243%	n.a.	n.a.	n.a.
up to 1 lag	0.211%	0.218%	0.278%	n.a.	n.a.
up to 2 lags	0.195%	0.212%	0.264%	0.323%	n.a.
up to 3 lags	0.187%	0.207%	0.238%	0.308%	0.361%
up to 4 lags	0.199%	0.221%	0.283%	0.324%	0.348%
up to 5 lags	0.227%	0.262%	0.322%	0.387%	0.354%

The baseline calibration of the model specifies a maximum number of three (quarterly) lags. The underlying presumption behind this choice is that it is unlikely that economic series contain reliable information about the future up to three-quarters ahead. It can be checked whether and how much the use of lags (introducing dynamics) actually improves the forecast accuracy. The dynamic structure of the model seems to improve the forecast accuracy (even for coincident estimations) and reduce confidence intervals. The forecast accuracy is enhanced by including up to three lags at all horizons. Conversely, the use of a fourth lag reduces the accuracy for coincident to three-quarter-ahead forecast. However, a fourth lag might be useful in the case of four-quarter-ahead forecasts[70].

Correlation of the coincident forecast index with GDP growth				
	1Q lag	Coincident	1Q lead	2Q lead
Coincident series (no lag)	66%	71%	56%	25%
up to 3 lags stacked series	83%	80%	61%	27%

Comparing the coincident forecast index with stacked lagged series to that with coincident series shows that only the RMSE is much lower with stacked lagged series (0.19% vs. 0.24%). The confidence intervals of bootstrapped forecasts are on average slightly wider, which is partially

[70] Less than half of the series (moving averages of survey or consumer price series) are available if only three lags are used. An observation for survey and CPI series is indeed available for the month following the coincident quarter.

due to less convergence with a smaller number of series[71]. But this feature is not systematic: the range for 2002Q1[72] is wider with coincident series, whereas ranges for 2002Q3 and Q4 are narrower (apparently the deceleration in growth was overestimated with lagged series).

The dynamic properties of the coincident forecast index are also modified with the use of stacked lagged series: its correlation with GDP is improved at all leads and lags. The index is coincident to GDP where predictors are used coincidently and slightly lagging based on stacked lagged series, but the coincident correlation is nevertheless considerably improved. PCA allows a more accurate extraction of coincident and leading common signals with the use of stacked lagged series. One might wonder whether there is a point in using VAR systems of factors[73] to retrieve the dynamics where PCA can actually directly do the job.

4.4. THE DATA COMPOSITION

Factor models do not involve just a handful of predictors but a large or very large number of them. In contrast to standard low-dimension regression models, measuring directly the exact contribution of each variable to forecasts would give meaningless results, given that there are too many of them (about 2000 in our empirical experiments) and, moreover, a one-to-one mapping between factors and predictors is not feasible.

Kapetanios and Marcellino (2003) introduce an interesting approach: they compute the impulse-response functions (IRF) of selected key indicators to some estimated factors. The IRF give some intuition about the nature of the shocks conveyed by the selected factors. However, the link to the underlying predictors is still missing and forecasts cannot be accounted for by the data themselves.

A more direct, albeit approximate, approach is possible in order to assess the linkage between predictors and forecasts. Since economic series are highly correlated within specific categories (e.g. survey, financial or employment series), it can make sense to check what the impact of a homogeneous subset of series of the same category is, instead of that of one particular predictor.We first examine the sensitivity of forecasts bootstrapped standard errors and point estimates to data composition in both forthcoming sections. A detailed analysis of the results obtained with an out-of-sample experiment covering 1999-2005 follows.

4.4.1. Is the pattern of survey series always consistent with that of other data? A first check with bootstrapped confidence intervals

The computation of standard errors around forecasts with cross-sectional bootstrap directly mirrors the degree of heterogeneity in the input data to the factor model. Suppose that a subset of series display a common pattern of strong growth whereas another subset points to stagnation, the common component to both subsets is estimated with more uncertainty than would be the case if all series exhibited the same pattern. The graph below shows that the standard error (SE) of quarterly

---

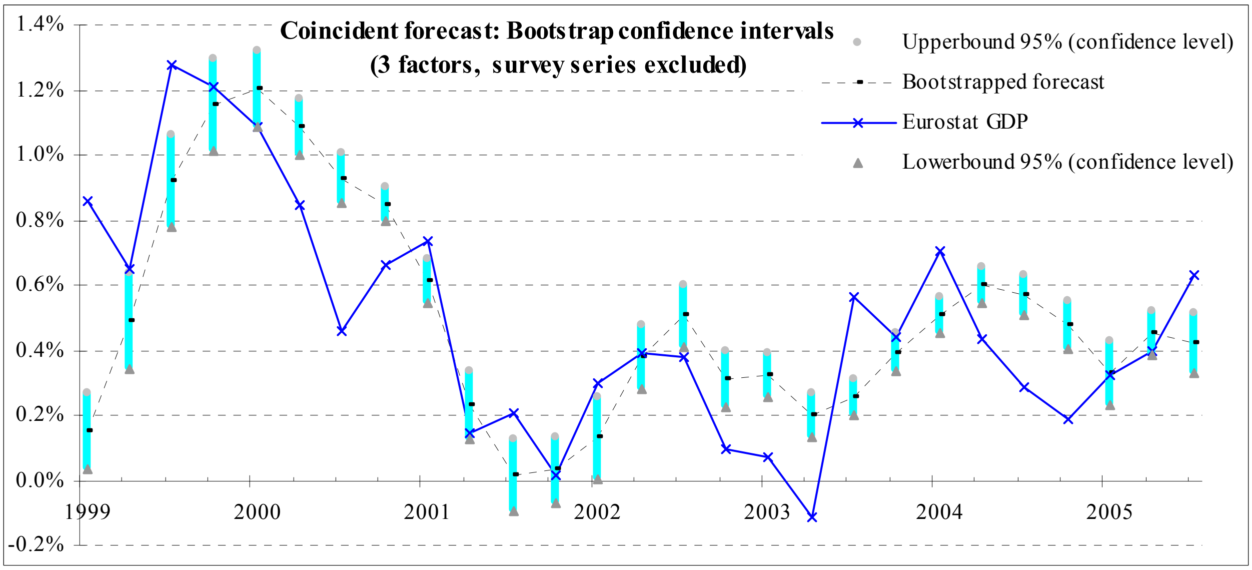
[71] The quadratic mean of the SE with all stacked series is 0.038% vs. 0.044% with coincident series only.

[72] And for 2005Q2, but given that the corresponding GDP estimate is likely to be substantially further revised, it is difficult to draw at present any conclusion from this quarter.

[73] Stock and Watson (1998) suggested the possibility of using stacked lagged series but the properties of common factors extracted from stacked series have not been examined until Grenouilleau (2004) to the best of our knowledge.

coincident[74] forecasts obtained with the factor model was much larger for very few specific quarters of the out-of-sample, for instance 2002Q3 to 2002Q4. The explanation mostly lies in the fact that diverging patterns were displayed by various subsets of data for these quarters (e.g. survey data vs. output or industrial production data). Then, it is more difficult to extract a common pattern summarised by a few factors.

A confirmation can be found in the bootstrapped forecasts obtained with data excluding survey series. The confidence intervals obtained for 2002Q3 and 2002Q4 with this data are narrower, which indicates that the pattern of survey series and that of other series were partially discrepant *for those quarters* and that the heterogeneity in data patterns is reduced by excluding survey series[75] in those specific cases.



The overall SE quadratic mean is nevertheless increased by 25% from 0.038% to 0.047% because of the loss of precision associated with a smaller number of predictors under PCA estimation. In general, it is nevertheless useful to use all series and not to discard survey series: the overall quadratic mean of standard errors *across all out-of-sample quarters* is increased by 25% from 0.038% to 0.047% where survey series are removed, because of the loss of precision associated with a smaller number of predictors (one third of the series are missing) with PCA estimation. This means that forecast uncertainty is generally increased where survey series are not used.

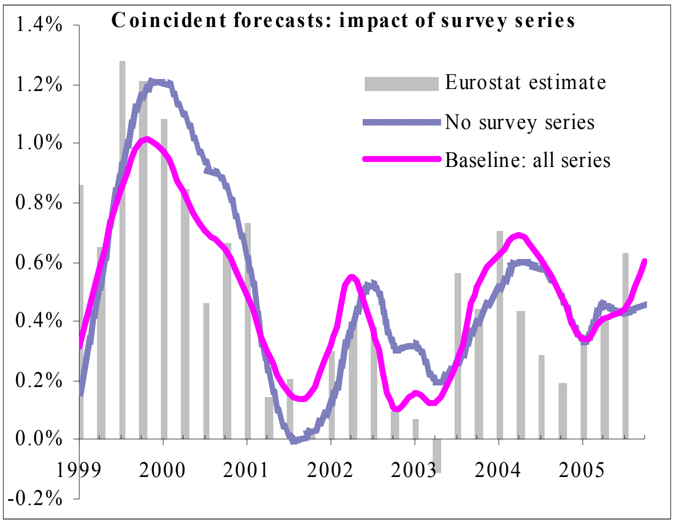
4.4.2. The contribution of various categories of data to point forecasts

In the previous paragraph, we examined the impact of the use of a specific category of series, survey series, on bootstrapped SE and confidence ranges. The impact of specific subsets of data on forecasts point estimates can also be monitored using the same method, i.e. comparing the forecast obtained with all series to that obtained with a specific subset of data removed from the dataset. The difference between both forecasts can be seen as a proxy for the marginal contribution of the latter subset of data to the forecast. The consistency of the comparison can be checked by using bootstrapped (unbiased) forecasts, which hardly differ from the sample forecasts, given that the dataset remains very large even without survey series.

[74] We call "coincident" the forecast performed for the current quarter as long as no GDP estimate is available (in practice, until 45 days after the end of the relevant quarter).

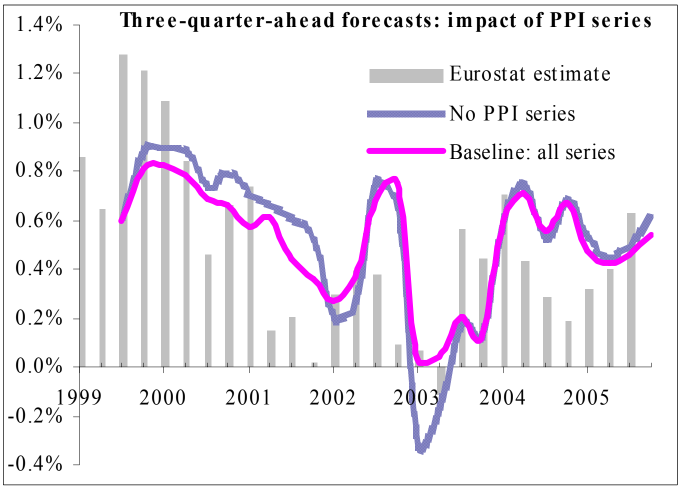
[75] In other words, there is a partial discrepancy in the patterns of survey series and other series in these specific quarters.

This marginal contribution should be seen as an indicator of the discrepancy between the common signal of the subset of removed series and that of the whole dataset. Where the difference between both forecasts is large, it means that although both common components are reflected by the same selected factors, their respective patterns differ strongly in the quarter of interest. For example, according to the graph above[76] survey data do not fully account for the buoyant growth recorded in 2000; the surge in real growth was most likely greater than the increase in confidence shown by surveys. Conversely, the decline in confidence following the Sept. 11th 2001 events was less sharp than the drop in economic activity consecutive to the fall of stock markets.



4.4.3. Some comments on the marginal contributions to the forecasts of selected categories of data

Some categories of data exhibit a larger contribution to forecasts than others[77]: this is the case of industrial confidence, consumer confidence, CPI, PPI at all forecast horizons. The contribution of industrial confidence varied substantially over the out-of-sample span; it was negative from 2000H2 onwards. The impact turned favourable during most 2002 in contrast to financial variables, which indicates that the slump on financial markets is not completely consistent with the economic situation in the industrial sector. Recently, the contribution of the industry survey was negative for most of 2005. In 2002H2 and early 2003, consumer confidence had a negative impact on forecasts (together with retail confidence). Consumer confidence exhibited a more subdued pattern than other series, which could reflect the fact that private consumption contributed to drag down economic growth.



CPI series generally conveyed a positive signal for the forecasts and particularly so for remote forecasts[78] over 1999-2001. CPI data also displayed a less subdued pattern than that of the real economy in 2002/03 for coincident or one-quarter-ahead forecasts. In both periods, forecasts have been relatively more optimistic based on consumer prices than based on all other data. The contribution of PPI series is also

[76] See graphs for other forecast horizons in annex 5.

[77] See detailed tables in annex 5.

[78] This is partially due to their timeliness, CPI series are a major subset for remote horizons forecasts (15% at the beginning of the out-of-sample span).

much more positive than that of other data in 2002H2 and 2003H1. Given their large contribution to more remote horizon forecasts, PPI series seem very well reflected by the three factors used in spite of the fact that their weight is less than 3% of the series (for three-quarter-ahead forecasts).

Two and three-quarter-ahead forecasts are particularly sensitive to construction confidence. The contribution to 2003H1 forecasts was very negative, which seems consistent with growth outturns. Service survey data have a very low contribution to forecasts, which could mean either that the pattern of services survey perfectly matches that of the whole economy (services are about 70% of the euro area economy) or that they contain a lot of noise or both[79]. Financial data usually display the same patterns as the whole dataset (especially at close forecast horizons). This result might be explained by the fact that stock market data usually incorporate all the information available plus a highly volatile component which is averaged out by PCA. Foreign trade data do not seem to convey much specific information over the out-of-sample span[80].

Employment and unemployment data show a small contribution to the forecasts but the effective contribution of the subset would be most likely greater, were the number of series collected larger. The sign of their contribution clearly lags the business cycle, being positive until 2002. It turns negative in 2003, suggesting that the deterioration of the job market situation weighed on the recovery. The contribution has decayed to nil for recent forecasts. Finally, the subsets of real data (industrial production or sales) do not exhibit significant contributions to the forecast. In the case of industrial production series, this is due to the fact that the common component to all series is likely to be close to that of output series. In the case of sales data, the subset is very small (between 1 to 2% of the dataset), which could partially explain the very small contribution of these series to the forecasts.

4.5. THE SORTING AND FILTERING OF INPUT SERIES ACCORDING TO A CRITERION OF CROSS-CORRELATION WITH THE SERIES TO BE FORECASTED

In Grenouilleau (2004), the model allows the filtering of series within a slightly modified EM algorithm. Series with cross-correlations to GDP below a certain threshold are removed from the input dataset. This intuition behind this feature is very simple: where predictors are too noisy, it could be better to remove them from the dataset in order to improve the extraction of the business cycle signal in the first few factors. A cross-correlation threshold of 0.2 yielded the best forecast accuracy in the latter paper. However, in the empirical applications of the present paper, the same threshold does not systematically enhance forecast accuracy with the new dataset used. Filtering with a threshold between 0.45 and 0.5 seems to improve the model accuracy more consistently; the RMSE is reduced by about 20% for the sub-sample 2001Q4-2003Q3, and by only 5% for the subsequent sub-sample 2003Q4-2005Q3 (the RMSE over this sub-sample is likely to change due to GDP revisions)[81]. Moreover, the model does not converge properly with a filtered dataset for out-of-sample quarters prior to 2001Q2.

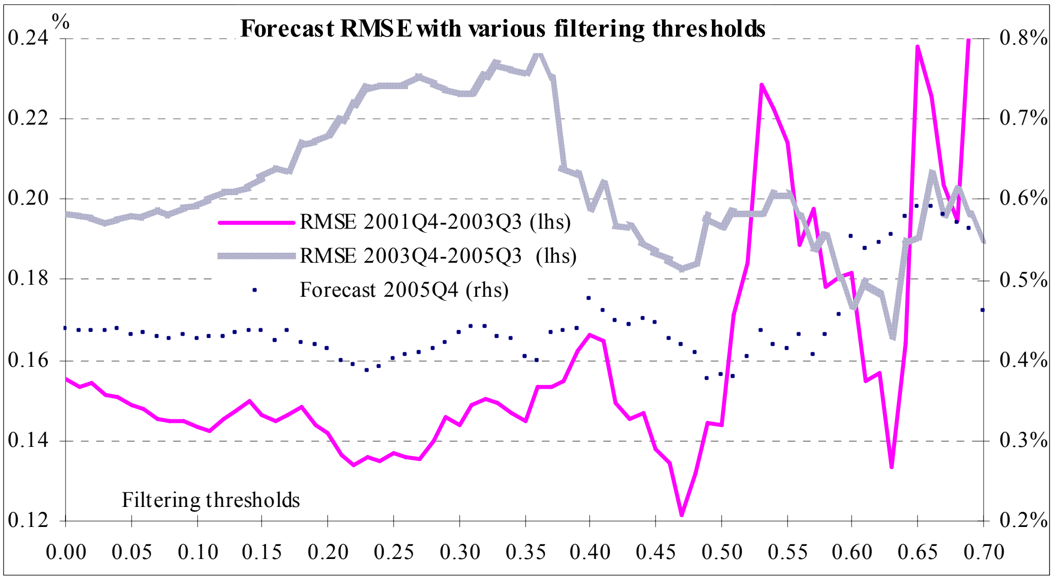
---

[79] Only the reading of loadings (for the selected factors) could give an answer: high loadings would mean that both patterns are usually similar, low loadings would mean that the subset series are considered as noise in the model.

[80] The fact that trade series are not timely available plays a certain role: most other series display an observation for the coincident quarter while this is not the case for the former (which are only used with lags in the model). The SE of bootstrapped forecasts is about twice as large where trade series are included. The RMSE is nevertheless slightly reduced from coincident to two-quarter-ahead forecasts, which indicate that these series are useful to extract a common component from the data although they are quite noisy.

[81] Bootstrap estimates are unfortunately not available due to computational resources constraint (the filtering of series is much more time consuming than standard algebraic operations with Matlab).

The removal of noise from the input series through filtering is limited by the fact that the GDP estimate available in real time contains a substantial share of noise compared to the final estimate. Actually, filtering according to the correlation with a noisy series available in real time might affect the signal extraction with PCA. An optimal filtering threshold might be associated with a given dataset, but it is difficult to estimate robustly in our empirical setup, given that the dataset composition changes across the out-of-sample period.



5. CONCLUSION

Approximate factor models offer a trade-off in terms of robustness to forecast uncertainties, which is completely different from standard small-size regression models. Using the classical regression typology adapted to factor models, future uncertainty can be strongly reduced with approximate factor models. Given the absence of restrictions on the number of potential predictors, future unknown shocks to the economy are likely to be better accounted for than with low-dimension systems, which have a tendency to describe past shocks well but do not necessarily capture present or future shocks. Cross-sectional bootstrap results also indicate that model uncertainty (here, uncertainty regarding the choice of predictors) is also dramatically reduced with our approximate factor model. This performance is partially due to the parsimony of the model specification. But the frontier for research in the field of approximate factor models remains the selection of factors. In our setup, the model's forecast accuracy is quite sensitive to the number of factors included. Information criteria based on regression fit (BIC) or bootstrapped eigenvalues (both averaged across out-of-sample quarters) suggest a factor selection which is consistent with out-of-sample results. Thus, some regularity in the factor structure is probably at play. But it might be possible to extract more information from the data with a fine-tuned selection of factors quarter by quarter.

The sensitivity analysis conducted in this paper suggests that the factor model forecast accuracy is structurally lower, within a range of "reasonable" calibration settings, than that of competing models, namely leading-indicator equations or AR models. The factor model's superior performance is particularly noticeable for forecasts performed at horizons between 2 months and 9 months ahead of GDP publication. Within this range, the approximate factor model forecasts are better behaved than those of indicator equations, in which the noise embodied in predictors rapidly increases parameter uncertainty and strongly reduces the forecast accuracy. At more remote horizons, standard stochastic models remain difficult to beat, probably because the signal in the reference series vanishes in the presence of noise and stochastic patterns due to time persistence and imperfect statistical adjustments. For coincident forecasting of GDP, factor models could also be more reliable than indicators equations, in which the fit typically deteriorates in the out-of-sample period after the calibration. The most striking feature is that our approximate factor model seems to deliver coincident forecasts that are as reliable as Flash or first estimates at the same date of release compared to final estimates. The factor model can thus provide an alternative coincident index of GDP growth with different statistical properties to quarterly National Accounts estimates but the same accuracy.

6. REFERENCES

Altissimo, F., A. Bassanetti, R. Cristadoro, M. Forni, M. Lippi, L. Reichlin, and G. Veronese (2001a). *A real time coincident indicator of the Euro area business cycle*, mimeo, Banca d'Italia (August).

Angelini E., J. Henry and R. Mestre (2001a). "A multi country trend indicator for euro area inflation: computation and properties". *ECB Working paper series* 60 (April).

Angelini E., J. Henry and R. Mestre (2001b). "Diffusion index-based inflation forecasts for the euro area". *ECB Working paper series* 61 (April).

Bai J. and S. Ng (2002). "Determining the number of factors in approximate factor models", *Econometrica* 70(1), 191-221.

Bai J. (2003). "Inferential theory for factor models of large dimensions", *Econometrica* 71(1), 135-171.

Bai J. and S. Ng (2005). "Determining the number of primitive shocks in factor models", *mimeo*, http://www-personal.umich.edu/~ngse/papers/dfac.pdf.

Banerjee A., M. Marcellino and I. Masten (2003). "Leading Indicators for Euro Area Inflation and GDP Growth", *CEPR Discussion paper* 3893 (May).

Beran (1985). "The Impact of the Bootstrap on Statistical Algorithms and Theory". *Statistical Science*, 18, No. 2, 175–184

Boivin J. and S. Ng (2003). "Are more data always better for factor analysis?" *NBER Working Paper* 9829 (July).

Boivin J. and S. Ng (2005). "Understanding and comparing factor-based methods", *mimeo,* http://www2.gsb.columbia.edu/faculty/JBoivin/mypapers/seattle.pdf.

Breitung J. and U. Kretschmer (2004). "Identification and estimation of dynamic factors from large macroeconomic panels", *mimeo*.

Camba-Mendez G., G. Kapetanios, R. J. Smith and M. R. Weale (1999). "An automatic leading indicator of European activity: Forecasting GDP growth for European countries", *mimeo*.

Cattell R. B. (1966). "The scree test for the number of factors". *Multivariate Behavioral Research* 1:245-76, 1966.

Chamberlain G. and M. Rothschild (1983). "Arbitrage and mean variance analysis on large asset markets", *Econometrica*, 51, 1281-1304.

Clements M. and D. Hendry (2005). "Evaluating a Model by Forecast Performance", *Oxford Bulletin of Economics and Statistics*, 67 (0/1), 931.

Connor G. and R. Korajczyk (1986). "Performance measurement with the arbitrage pricing theory", *Journal of Financial Economics*, 15, 373-394.

Connor G. and R. Korajczyk (1993). "A test for the number of factors in an approximate factor model", *The Journal of Finance*, Vol. 48, No. 4, 1263-1291.

Daudin J. J., C. Duby and P. Trecourt (1988). "Stability of Principal Component Analysis Studied by the Bootstrap Method", *Statistics*, 19-2, 241-258.

Diebold F. X. and R. S. Mariano (1995). "Comparing predictive accuracy", *Journal of Business and Economic Statistics*, Vol. 13, 253-263, 3 (July).

Doz C. and F. Lenglart (1999). Analyse factorielle dynamique : test du nombre de facteurs, estimation et application à l'enquête de conjoncture dans l'industrie. *Annales d'économie et de statistique*, 1999/54.

Efron B. and J. Tibshirani (1993). *An Introduction to the Bootstrap*. Chapman & Hall, New York.

Erkel-Rousse, H. and G. Prioux (2002). L'apport des enquêtes de conjoncture dans les différents secteurs d'activité à l'analyse conjoncturelle. *Note de conjoncture*, INSEE (June).

Eurostat (2003). "Flash estimates of the European quarterly GDP: the approach of Eurostat", *mimeo*.

Forni M., M. Hallin, M. Lippi and L. Reichlin (2001), "Coincident and leading indicators for the Euro area", *The Economic Journal* 111, 62-85.

Forni M., M. Hallin, M. Lippi and L. Reichlin (2000) "The generalized dynamic factor model: identification and estimation", *The Review of Economics and Statistics* (November).

Garratt A., Lee K., Pesaran M.H. and Shin Y. (2003). "Forecast uncertainties in macroeconometric modeling: an application to the UK economy". *Journal of the American Statistical Association*, 98, 829-838.

Geweke J. (1977). "Labor turnover and employment dynamics in US manufacturing industry" in *New methods in business cycle research*, ed. C. Sims. Federal Reserve Bank of Minneapolis.

Grenouilleau D. (2004). "A sorted leading indicators dynamic (SLID) factor model for short-run euro-area GDP forecasting ". *European Economy - Economic Paper* n°219.

Hallin M. and R. Liska (2005). "The Generalized Dynamic Factor Model: Determining the number of factors", *mimeo*.

Härdle W., Horowitz J. and Kreiss J.P. (2003) "A Review of the Bootstrap for Time Series". *International Statistical Review*, 71, 435-459.

Harvey D., S. Leybourne and P. Newbold (1997). "Testing the equality of prediction mean squared errors", *International Journal of Forecasting* 13, 281-291.

Kapetanios G. and M. Marcellino (2003). "Structural impulse response analysis of dynamic factor models for large datasets", *mimeo*.

Lawley D. N. and A. E. Maxwell (1971). *Factor Analysis as a Statistical Method*. Macmillan, New-York.

Perron P. (1989). "The great crash, the oil price shock, and the unit root hypothesis", *Econometrica*, 57(6), 1361-1401.

Quah D. and T. Sargent (1993). "A dynamic index model for large cross sections" in *Business cycles, indicators, and forecasting*, ed. Stock J.H. and M.W. Watson. NBER research studies in business cycle. University of Chicago press.

Reichlin L. (2002). "Factor models in large cross-sections of time series", *mimeo*.

Sargent T. and C. Sims (1997). "Business cycle modelling without pretending to have too much a priori economic theory" in *New methods in business cycle research*, ed. C. Sims. Federal Reserve Bank of Minneapolis.

Sédillot F. and N. Pain (2003). "Indicator models of real GDP growth in selected OECD countries", *OECD Economic Department working papers* 364.

Stock J.H. and M.W. Watson (1988). "A probability model of the coincident economic indicators", *NBER Working Paper* 2772 (November).

Stock J.H. and M.W. Watson (1989). "New indexes of coincident and leading economic indicators", *NBER Macroeconomic Annual,* 351-394.

Stock J.H. and M.W. Watson (1998). "Diffusion indexes". *NBER Working Paper* 6702 (August).

Stock J.H. and M.W. Watson (1999). "Forecasting inflation". *NBER Working Paper* 7023 (March).

Stock J.H. and M.W. Watson (2002a). "Macroeconomic forecasting using diffusion indexes". *Journal of Business and Economic Statistics* 20, 147-162.

Stock J.H. and M.W. Watson (2002b). "Forecasting using principal components from a large number of predictors". *Journal of the American statistical association* 97 (460), 1167-1179.

Stock J.H. and M.W. Watson (2003). "Forecasting output and inflation: the role of asset prices". *Journal of Economic Literature* Vol. XLI, 788-829, Sept.

Stock J.H. and M.W. Watson (2003). *Introduction to econometrics*. Addison-Wesley, Boston.

Stock J.H. and M.W. Watson (2004). "Forecasting with many predictors", survey prepared for the forthcoming *Handbook of Economic Forecasting*.

7. ANNEXES

ANNEX 1: EM ALGORITHM USED IN THE SLID FACTOR MODEL[82]

It is possible to estimate f^h and Λ^h with principal component analysis in the following equation (2') for all N_h series and all T+h time observations, if and only if X^h does not contain any missing observations. By definition, all series in X^h (except the last series GDP_{T+h}) do not display any missing observations (those which did were removed from the dataset). The matrix X^h is never empty irrespective of the forecast horizons h, since it is always possible to stack the predictors with additional lags. At each forecast horizon h, GDP_{T+h} displays only one missing observation, the last observation T+h, since it is possible to use recursively forecasts obtained at previous horizons h_j for j<h.

$$X^h = f^h \cdot \Lambda^h + U^h \tag{1'}$$

$$\mathrm{GDP}_{\mathrm{T+h}} = X^h_{T+h,N_h} \tag{2'}$$

The EM algorithm consists in three steps:

(a) Plug in a first guess for the forecast (3'),

(b) Obtain joint estimates of f^h and Λ^h,

(c) Re-estimate the forecast GDP_{T+h} based on an exogenous projection of GDP onto a few eigenvectors assumed to be latent factors for GDP,

(d) Iterate sequentially steps (b) and (c) until convergence.

There are several ways to technically carry out steps (a) and (c). The estimation of GDP_{T+h} at step (c) can be derived from a projection of GDP onto a set of relevant factors estimated with OLS or a VAR system. We suggest a more simple[83] solution (4'), which consists of approximating the regression coefficients with the loadings of GDP corresponding to the various factors.

$$\mathrm{GDP}_{\mathrm{T+h}} = \sum_{i \in I_f} f^h_{T+h,i} . \Lambda^h_{i,T+h} \ , \tag{3'}$$

where I_f is the set of indexes corresponding to the latent factors for GDP.

To initialise the iterations at step (a), one can either replace the missing observation by a first guess obtained using the same technique as step (c) but with a dataset trimmed from the last observation (equation 5'), or one can plug in a value of 0. The final result is the same and the computation time needed about the same (the algorithm usually converges rapidly) in either case.

$$\mathrm{GDP}_{\mathrm{T+h}} = \sum_{i \in I_f} f^{h-1}_{T+h-1,i} . \Lambda^{h-1}_{i,T+h-1} \quad \text{(with the same definition for } I_f) \tag{4'}$$

---

[82] In Grenouilleau (2004), a filtering of input series according to their cross-correlation with GDP is performed prior to the extraction of principal components.

[83] In principle, the results should be identical to those from an OLS regression

ANNEX 2: STATISTICS OF FORECAST ACCURACY

The statistics of model comparison are the forecasts RMSE and the Diebold-Mariano (1995) (DM) statistics of compared predictive accuracy. It is assumed that the SLID factor model and the benchmark models are not nested, insofar as factor estimates do not result from a trivial linear combination of variables (possibly lagged) used in other models. The DM statistic is used in its most simple form:

$\sqrt{n}\cdot\frac{\bar{d}}{\hat{\sigma}_d}\xrightarrow{d}\mathrm{N}(0,1)$, where: n is the number of observations of the experiment,

$\bar{d}=\frac{1}{n}\sum_{t=1}^{n}d_t$, $d_t=e_{at}^2-e_{bt}^2$, e_{at}^2, e_{bt}^2 are the respective squared forecast errors of models a and b, and $\hat{\sigma}_d^2=\frac{1}{n}\sum_{t=1}^{n}(d_t-\bar{d})^2$.

A slightly adapted version of the Diebold-Mariano statistics (DM) can be used for an unbiased empirical variance in the case of finite samples:

$\sqrt{n-1}\cdot\frac{\bar{d}}{\hat{\sigma}'_d}\rightarrow St(n-1)$ where $\hat{\sigma}'^2_d=\frac{1}{n-1}\sum_{t-1}^{n}(d_t-\bar{d})^2$.

A heteroskedasticity and autocorrelation consistent (HAC) estimate of the empirical variance using a Bartlett lag window (as autocovariance weighting scheme[84]) was computed. In addition, the correction for small sample suggested by Harvey *et al.* [1997] is introduced in the Student statistics:

$$\sqrt{n+1-2h+\frac{h(h-1)}{n}}\cdot\frac{\bar{d}}{\hat{\sigma}''_d}\rightarrow St(n-1), \tag{5'}$$

where $$\hat{\sigma}''^2_d=\frac{1}{n}\sum_{t=1}^{n}(d_t-\bar{d})^2+\frac{2}{n}\sum_{j=1}^{2}(1-\frac{j}{3})\sum_{t=1}^{n-j}(d_t-\bar{d})\cdot(d_{t+j}-\bar{d}) \tag{6'}$$

The test's null hypothesis is: models a and b do not exhibit any difference in forecast accuracy, against the alternative hypothesis: model a is superior to model b (if $\bar{d}<0$) or model b is superior to model a (if $\bar{d}>0$). Models are compared at a confidence level of 95%.

[84] See Newey and West [1987] for the properties of such an estimator. Autocovariances up to the order 2 (truncation parameter) are included in the variance estimator. The econometric literature does not provide a universal criterion for the choice of the truncation parameter and the relevant growth rate of this parameter with n. Newey and West introduce results for a slower than $n^{1/4}$ growth rate. Stock and Watson [2003] suggest $0.75n^{1/3}$. Both rules suggest a choice of 2.

ANNEX 3: OUT-OF-SAMPLE RESULTS

3.1. Statistics of compared accuracy

◊ *Coincident forecast compared to vintage Eurostat estimates*

Flash/first est.		Forecast errors				Forecast of GDP revision direction		
GDP		SLID coincident	Latest Eurostat est.	Eurostat first/flash est.	SLID coincident	all quarters	excl. differences < 0.05%	excl. differences < 0.1%
1999Q1	43	0.31%	0.85%	0.40%	0.31%	0	0	
1999Q2	44	0.57%	0.65%	0.30%	0.57%	1	1	1
1999Q3	45	0.85%	1.27%	1.00%	0.85%	0	0	0
1999Q4	46	1.01%	1.20%	0.90%	1.01%	1	1	1
2000Q1	47	0.97%	1.08%	0.70%	0.97%	1	1	1
2000Q2	48	0.84%	0.84%	0.90%	0.84%	1	1	
2000Q3	49	0.70%	0.46%	0.70%	0.70%	0		
2000Q4	50	0.65%	0.66%	0.70%	0.65%	1	1	
2001Q1	51	0.49%	0.73%	0.50%	0.49%	0		
2001Q2	52	0.30%	0.15%	0.10%	0.30%	1	1	1
2001Q3	53	0.15%	0.21%	0.10%	0.15%	1	1	
2001Q4	54	0.15%	0.02%	-0.20%	0.15%	1	1	1
2002Q1	55	0.32%	0.30%	0.22%	0.32%	1	1	1
2002Q2	56	0.55%	0.39%	0.34%	0.55%	1	1	1
2002Q3	57	0.37%	0.38%	0.33%	0.37%	1		
2002Q4	58	0.11%	0.09%	0.17%	0.11%	1	1	
2003Q1	59	0.16%	0.07%	0.01%	0.16%	1	1	1
2003Q2	60	0.12%	-0.11%	-0.08%	0.12%	0	0	0
2003Q3	61	0.26%	0.56%	0.38%	0.26%	0	0	0
2003Q4	62	0.52%	0.44%	0.31%	0.52%	1	1	1
2004Q1	63	0.63%	0.70%	0.57%	0.63%	1	1	
2004Q2	64	0.69%	0.44%	0.51%	0.69%	0	0	0
2004Q3	65	0.61%	0.28%	0.30%	0.61%	0	0	0
2004Q4	66	0.49%	0.19%	0.15%	0.49%	1	1	1
2005Q1	67	0.34%	0.32%	0.50%	0.34%	1	1	1
2005Q2	68	0.41%	0.40%	0.29%	0.41%	1	1	1
2005Q3	69	0.44%	0.63%	0.64%	0.44%	1	1	1
RMSE 99Q1-04Q2		0.207%		0.203%	0.162%	70%	75%	72%
compared to…		latest est.		latest est.	first/flash est			
Range at 68% confidence level		±0.2		±0.2	±0.2			
Bias 99Q1-04Q2		-0.03%		-0.11%	0.08%			
Diebold-Mariano			benchmark:	first/flash est.				
99Q1-04Q2				58%				
				0.20				

◊ *Coincident forecast compared to Eurocoin (revised index)*

Eurocoin		Forecast errors							Forecast of GDP revision direction*		
GDP		Latest Eurostat est.	Eurocoin coincident	SLID coincident	Eurostat first/flash est.	SLID 1Q ahead	Eurocoin 2M ahead	Eurocoin 1Q ahead	all quarters	excl. differences < 0.05%	excl. differences < 0.1%
1999Q1	43	0.85%	0.42%	0.31%	0.40%				1		
1999Q2	44	0.65%	0.65%	0.57%	0.30%	0.51%	0.48%	0.42%	1	1	1
1999Q3	45	1.27%	1.01%	0.85%	1.00%	0.79%	0.76%	0.65%	1		
1999Q4	46	1.20%	1.27%	1.01%	0.90%	0.99%	1.14%	1.01%	1	1	1
2000Q1	47	1.08%	1.08%	0.97%	0.70%	0.94%	1.23%	1.27%	1	1	1
2000Q2	48	0.84%	1.00%	0.84%	0.90%	0.79%	1.04%	1.08%	0	0	
2000Q3	49	0.46%	0.79%	0.70%	0.70%	0.70%	0.96%	1.00%	0	0	
2000Q4	50	0.66%	0.59%	0.65%	0.70%	0.59%	0.71%	0.79%	1	1	1
2001Q1	51	0.73%	0.40%	0.49%	0.50%	0.56%	0.52%	0.59%	0	0	
2001Q2	52	0.15%	0.21%	0.30%	0.10%	0.34%	0.34%	0.40%	1	1	1
2001Q3	53	0.21%	0.06%	0.15%	0.10%	0.18%	0.15%	0.21%	0		
2001Q4	54	0.02%	-0.12%	0.15%	-0.20%	0.21%	-0.01%	0.06%	1	1	
2002Q1	55	0.30%	0.49%	0.32%	0.22%	0.27%	0.02%	-0.12%	1	1	1
2002Q2	56	0.39%	0.56%	0.55%	0.34%	0.54%	0.59%	0.49%	1	1	1
2002Q3	57	0.38%	0.40%	0.37%	0.33%	0.41%	0.50%	0.56%	1	1	
2002Q4	58	0.09%	0.40%	0.11%	0.17%	0.09%	0.42%	0.40%	0	0	0
2003Q1	59	0.07%	0.40%	0.16%	0.01%	0.12%	0.35%	0.40%	1	1	1
2003Q2	60	-0.11%	0.43%	0.12%	-0.08%	0.20%	0.42%	0.40%	0	0	0
2003Q3	61	0.56%	0.59%	0.26%	0.38%	0.19%	0.48%	0.43%	1	1	1
2003Q4	62	0.44%	0.70%	0.52%	0.31%	0.50%	0.64%	0.59%	1	1	1
2004Q1	63	0.70%	0.54%	0.63%	0.57%	0.66%	0.65%	0.70%	0		
2004Q2	64	0.44%	0.58%	0.69%	0.51%	0.68%	0.53%	0.54%	0	0	
2004Q3	65	0.28%	0.48%	0.61%	0.30%	0.65%	0.57%	0.58%	0	0	0
2004Q4	66	0.19%	0.39%	0.49%	0.15%	0.53%	0.44%	0.48%	1	1	1
2005Q1	67	0.32%	0.36%	0.34%	0.50%	0.42%	0.38%	0.39%	1	1	1
2005Q2	68	0.40%	0.35%	0.41%	0.29%	0.40%	0.36%	0.36%	1	1	
2005Q3	69	0.63%	0.40%	0.44%	0.64%	0.46%	0.36%	0.35%	1	1	1
RMSE 99Q3-05Q3			0.216%	0.187%		0.208%	0.248%	0.275%	67%	70%	81%
compared to…			latest est.	latest est.		latest est.	latest est.	latest est.	*Eurocoin (coincident)		
Range at 68% confidence level			±0.2	±0.2		±0.2	±0.2	±0.3			
Bias 99Q3-05Q3			0.07%	0.02%	-0.07%	0.02%	0.07%	0.08%			
Diebold-Mariano			benchmark:	SLID coin.	flash/first est.	benchmark:	SLID 1Q ahead	SLID 1Q ahead			
99Q3-05Q3				77%	83%		90%	97%			
				0.75	0.96		1.30	2.05			

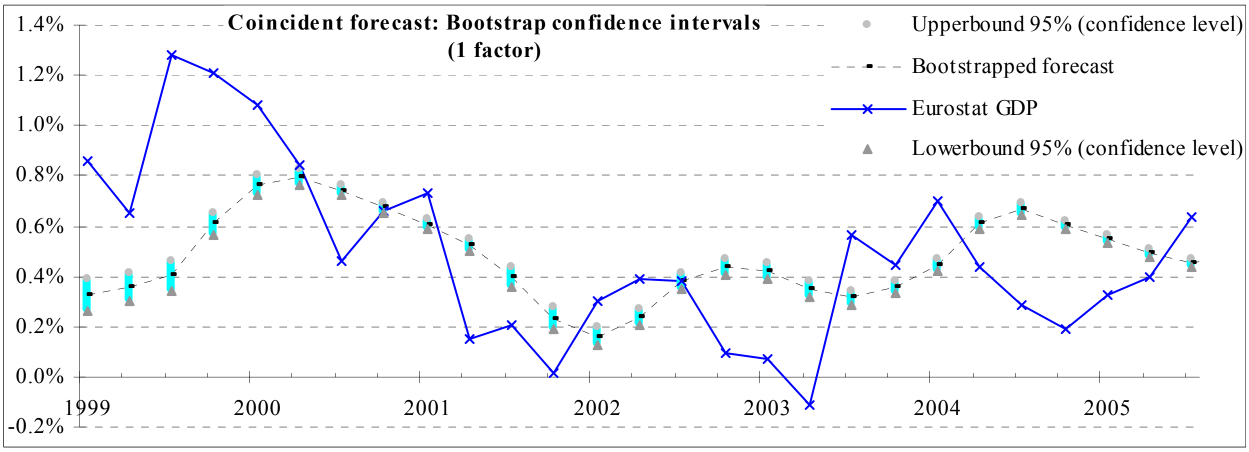
◊ *All horizons forecasts compared to AR models*

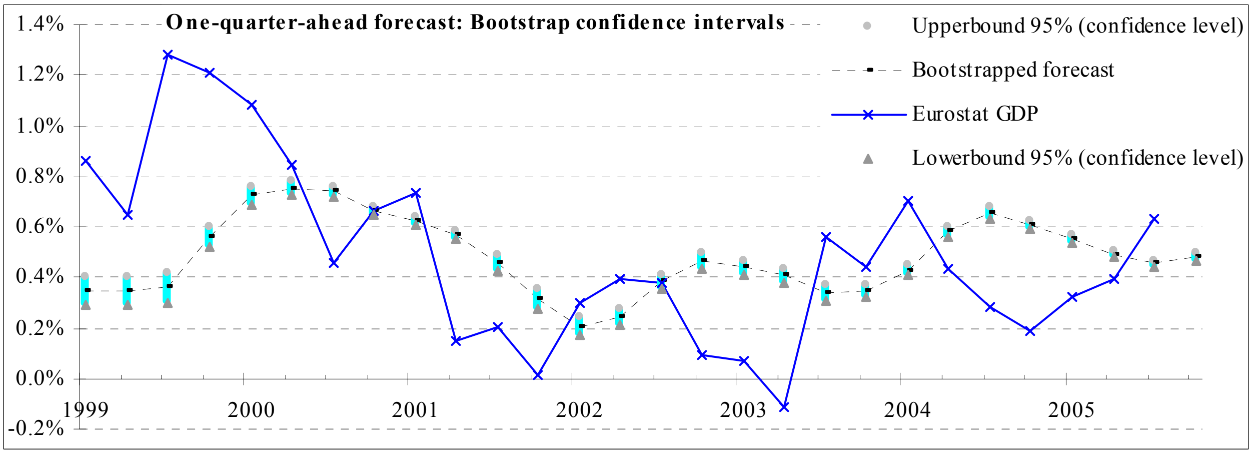
AR model		Forecast errors						
GDP		Latest Eurostat est.	SLID 1Q ahead	SLID 2Q ahead	SLID 3Q ahead	AR(1)	AR(2)	AR(3)
1999Q1	43	0.85%	0.24%			0.33%		
1999Q2	44	0.65%	0.51%	0.37%		0.61%	0.42%	
1999Q3	45	1.27%	0.79%	0.72%	0.62%	0.54%	0.52%	0.40%
1999Q4	46	1.20%	0.99%	0.90%	0.83%	0.82%	0.50%	0.51%
2000Q1	47	1.08%	0.94%	0.92%	0.85%	0.83%	0.63%	0.49%
2000Q2	48	0.84%	0.79%	0.83%	0.79%	0.79%	0.65%	0.58%
2000Q3	49	0.46%	0.70%	0.66%	0.69%	0.69%	0.67%	0.62%
2000Q4	50	0.66%	0.59%	0.62%	0.68%	0.50%	0.62%	0.63%
2001Q1	51	0.73%	0.56%	0.55%	0.58%	0.60%	0.53%	0.60%
2001Q2	52	0.15%	0.34%	0.52%	0.67%	0.63%	0.58%	0.53%
2001Q3	53	0.21%	0.18%	0.20%	0.43%	0.36%	0.60%	0.56%
2001Q4	54	0.02%	0.21%	0.27%	0.37%	0.38%	0.46%	0.58%
2002Q1	55	0.30%	0.27%	0.31%	0.27%	0.28%	0.47%	0.46%
2002Q2	56	0.39%	0.54%	0.44%	0.37%	0.41%	0.41%	0.47%
2002Q3	57	0.38%	0.41%	0.65%	0.68%	0.45%	0.47%	0.43%
2002Q4	58	0.09%	0.09%	0.20%	0.74%	0.45%	0.49%	0.47%
2003Q1	59	0.07%	0.12%	0.02%	0.02%	0.30%	0.48%	0.48%
2003Q2	60	-0.11%	0.20%	0.18%	0.03%	0.28%	0.41%	0.48%
2003Q3	61	0.56%	0.19%	0.19%	0.19%	0.18%	0.39%	0.42%
2003Q4	62	0.44%	0.50%	0.31%	0.11%	0.52%	0.33%	0.40%
2004Q1	63	0.70%	0.66%	0.67%	0.61%	0.46%	0.50%	0.35%
2004Q2	64	0.44%	0.68%	0.68%	0.72%	0.59%	0.47%	0.48%
2004Q3	65	0.28%	0.65%	0.64%	0.56%	0.46%	0.54%	0.46%
2004Q4	66	0.19%	0.53%	0.60%	0.68%	0.38%	0.48%	0.51%
2005Q1	67	0.32%	0.42%	0.43%	0.48%	0.34%	0.44%	0.47%
2005Q2	68	0.40%	0.40%	0.41%	0.43%	0.40%	0.41%	0.44%
2005Q3	69	0.63%	0.46%	0.46%	0.46%	0.43%	0.44%	0.42%
RMSE 99Q3-05Q3			0.208%	0.238%	0.308%	0.276%	0.337%	0.364%
compared to…			latest est.	latest est.	latest est.	latest est.	latest est.	latest est.
Range at 68% confidence level			±0.2	±0.2	±0.3	±0.3	±0.3	±0.4
Bias 99Q3-05Q3			0.02%	0.03%	0.05%	0.01%	0.03%	0.02%
Diebold-Mariano		benchmark:	AR(1)	AR(2)	AR(3)			
99Q3-05Q3			95%	95%	82%			
			1.67	1.71	0.92			

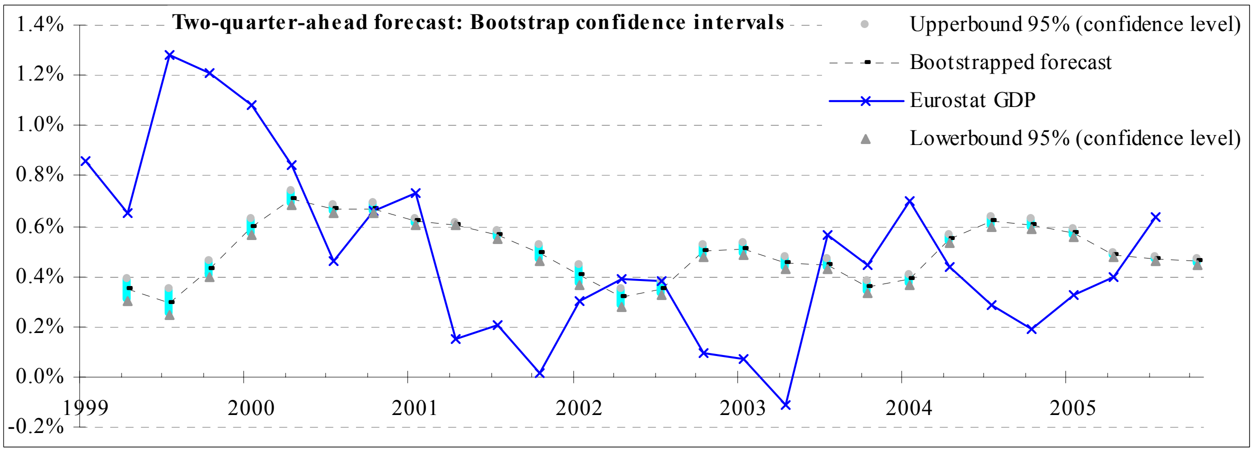
3.2. Bootstrapped forecasts

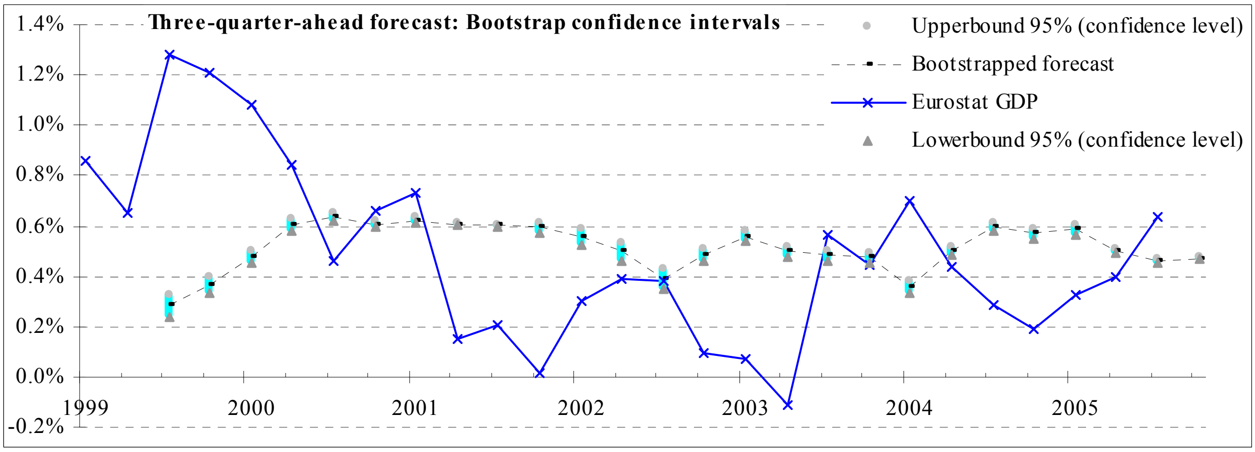
The following graphs display bootstrapped forecasts confidence intervals for all out-of-sample quarters. The same graphs are produced for various factor combinations.

◊ *Forecasts confidence intervals with one factor*

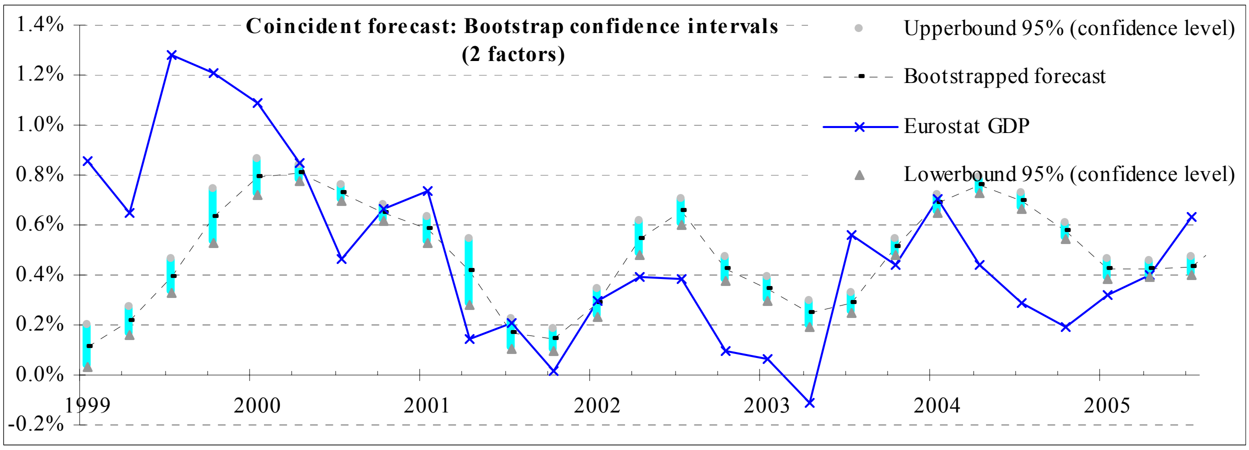


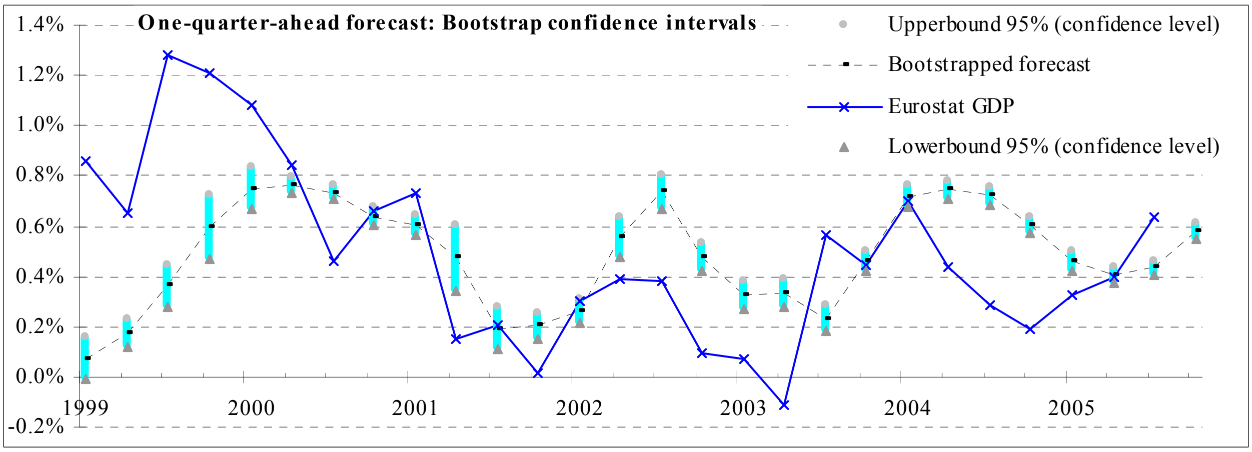


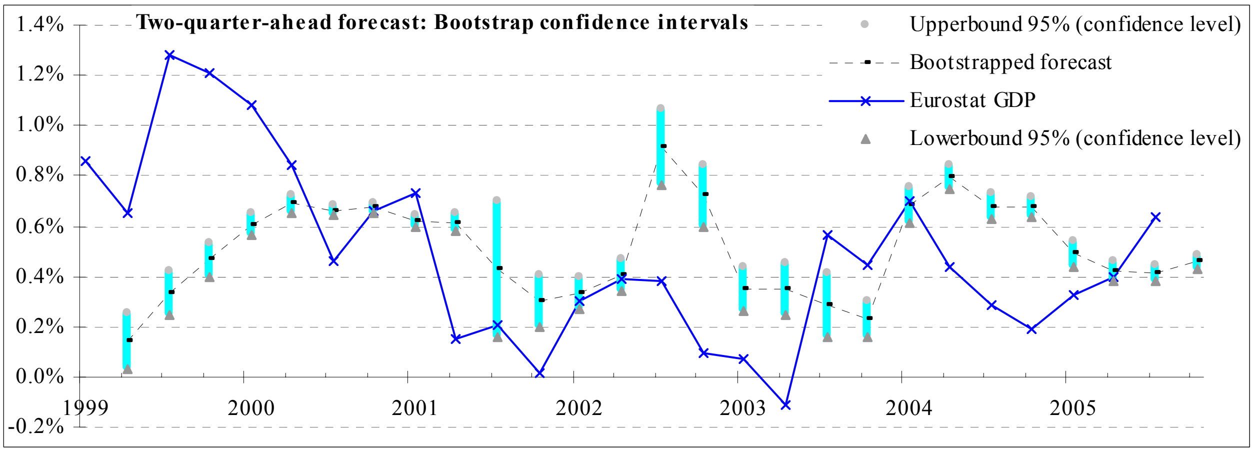


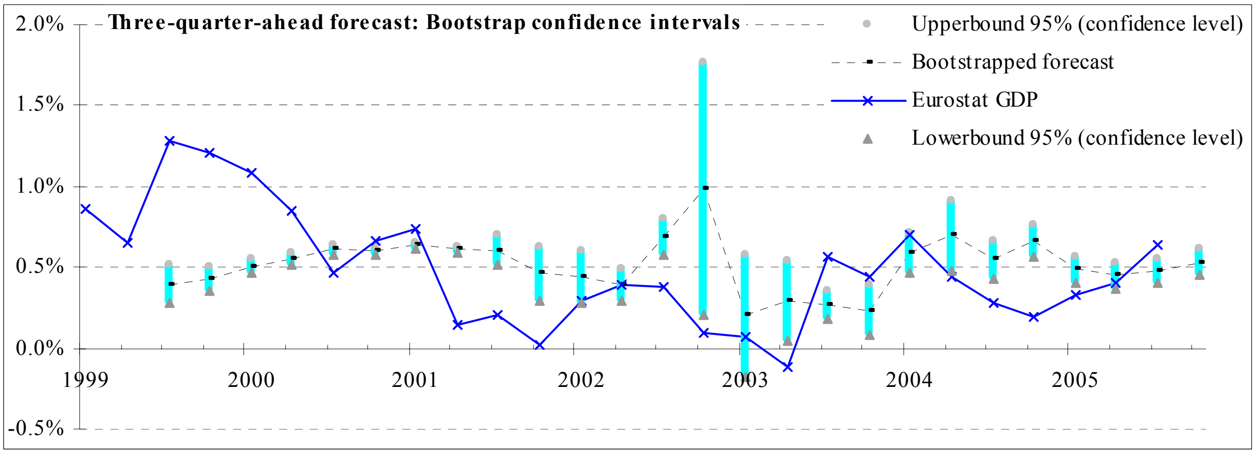


◊ *Forecasts confidence intervals with two factors*

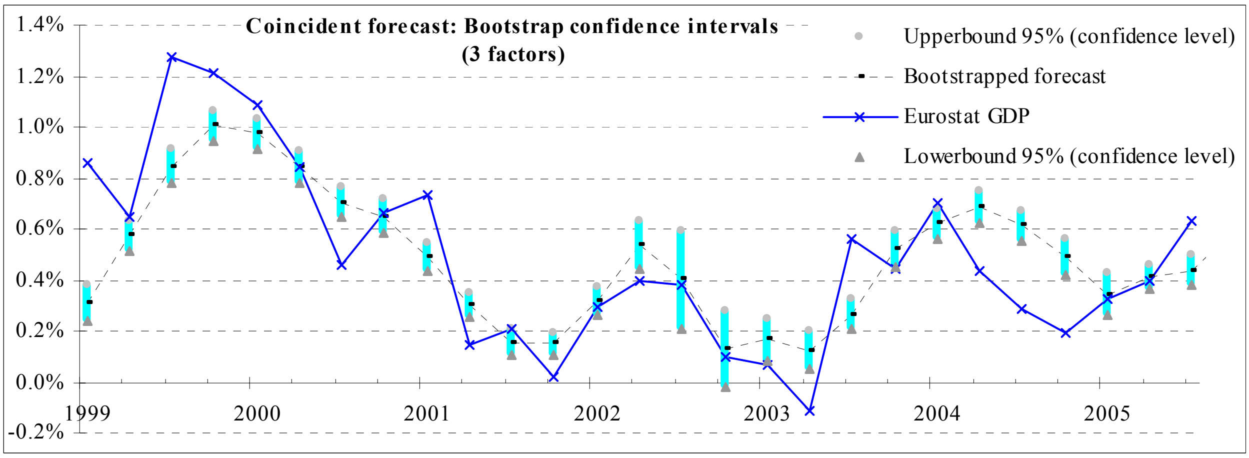


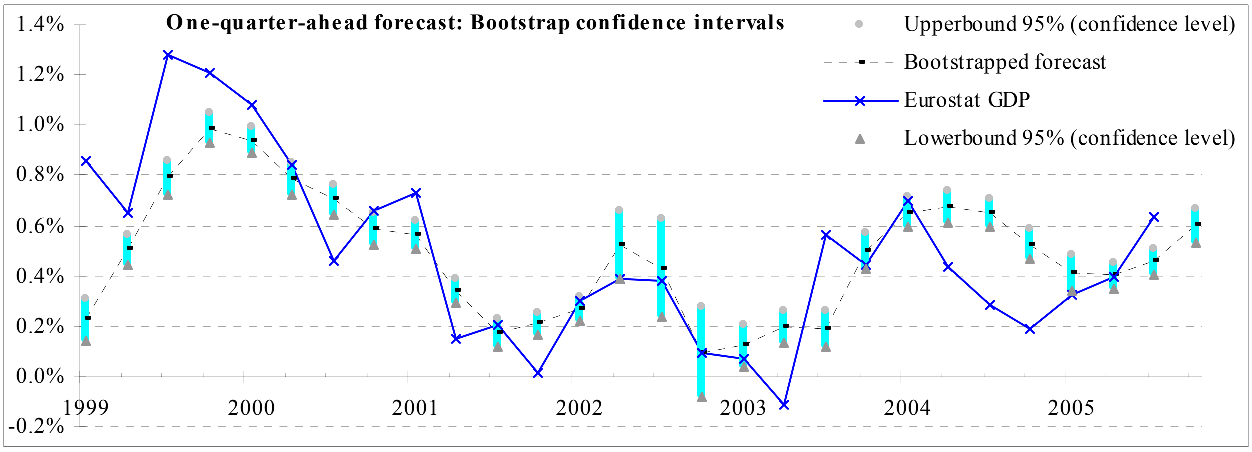


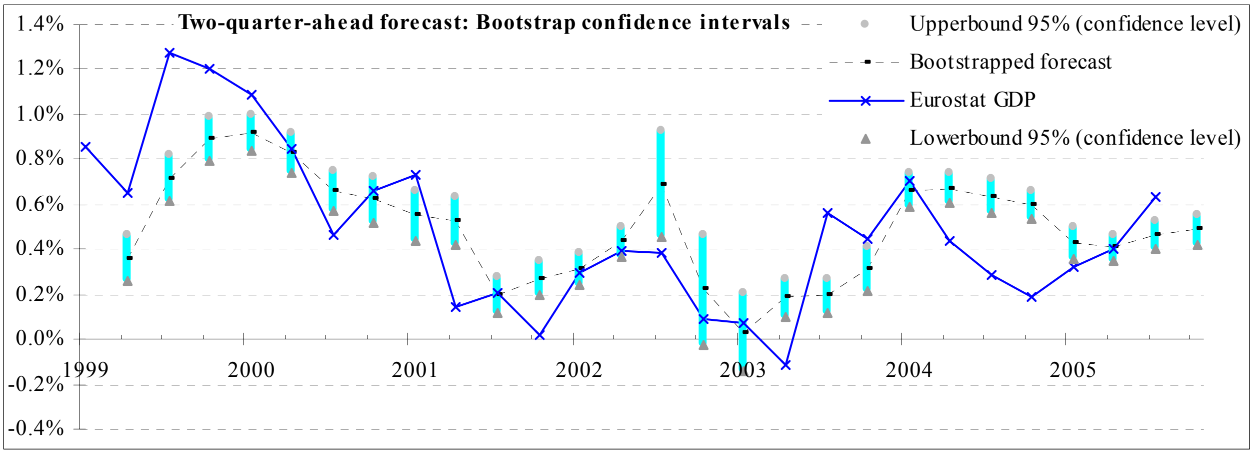


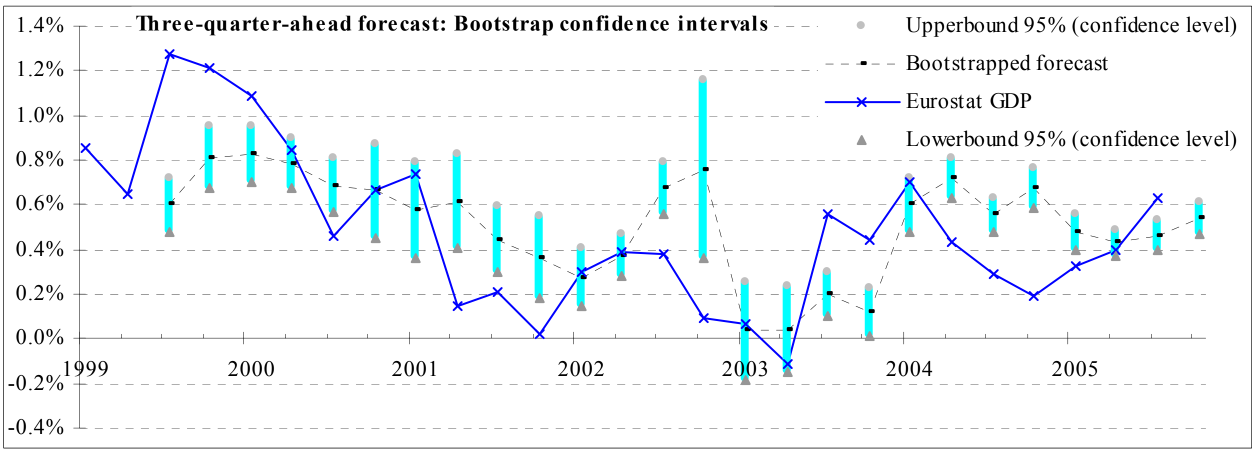


◊ *Forecasts confidence intervals with three factors*

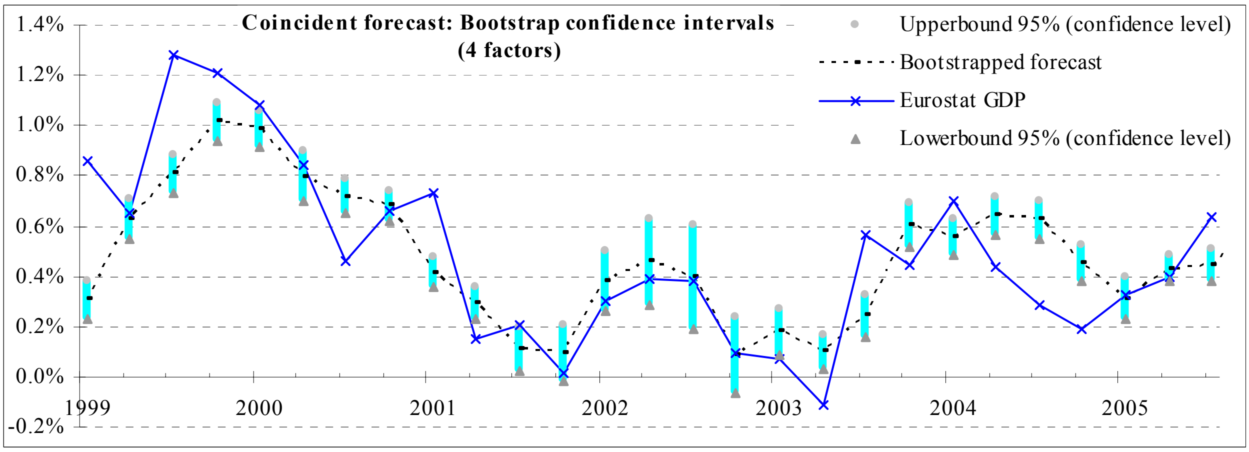


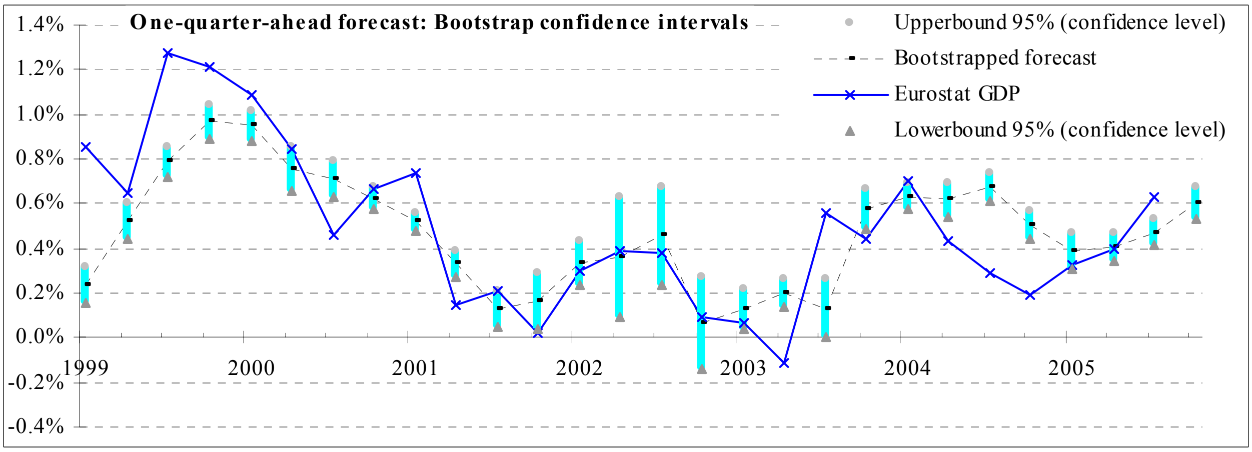


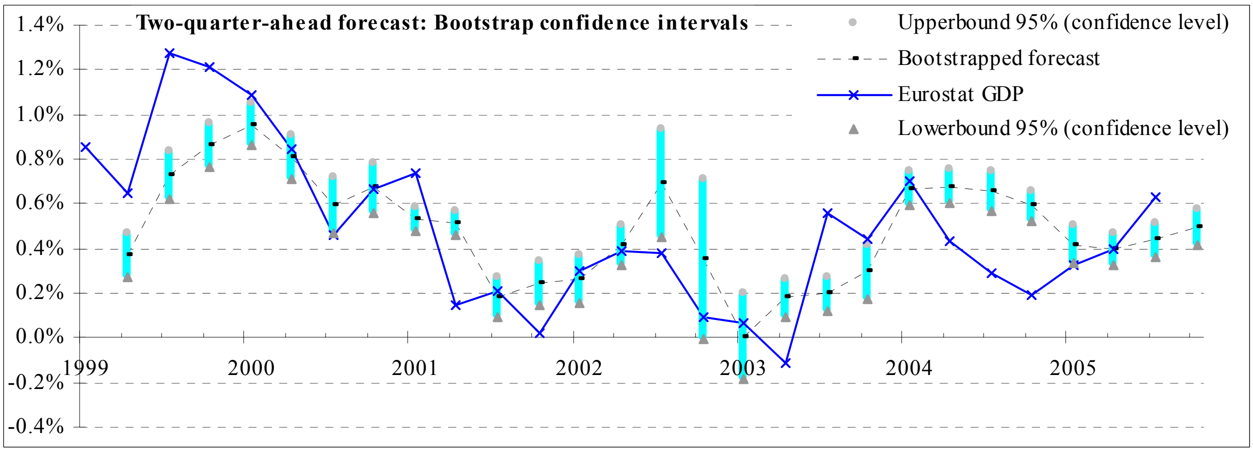


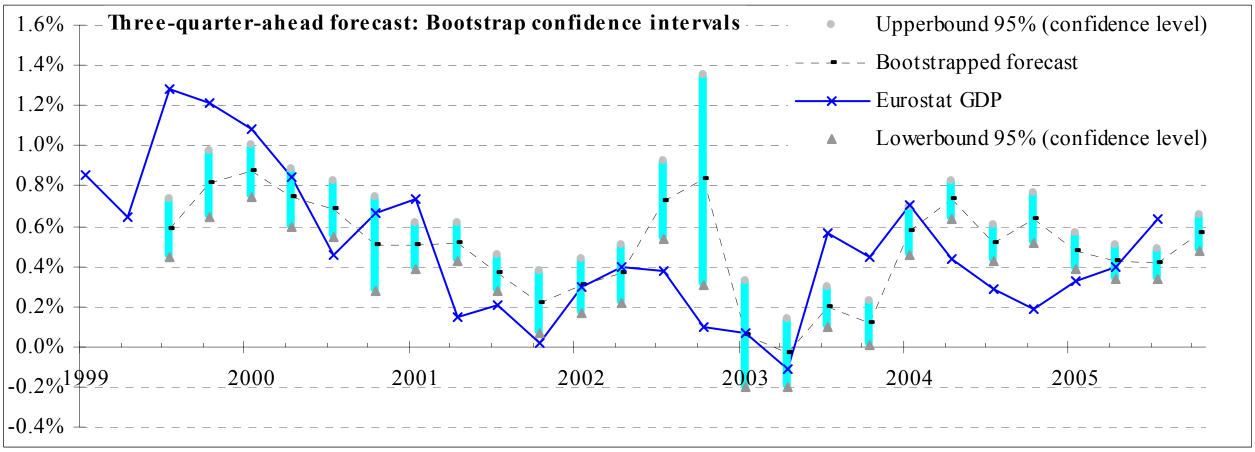


◊ *Forecasts confidence intervals with four factors*

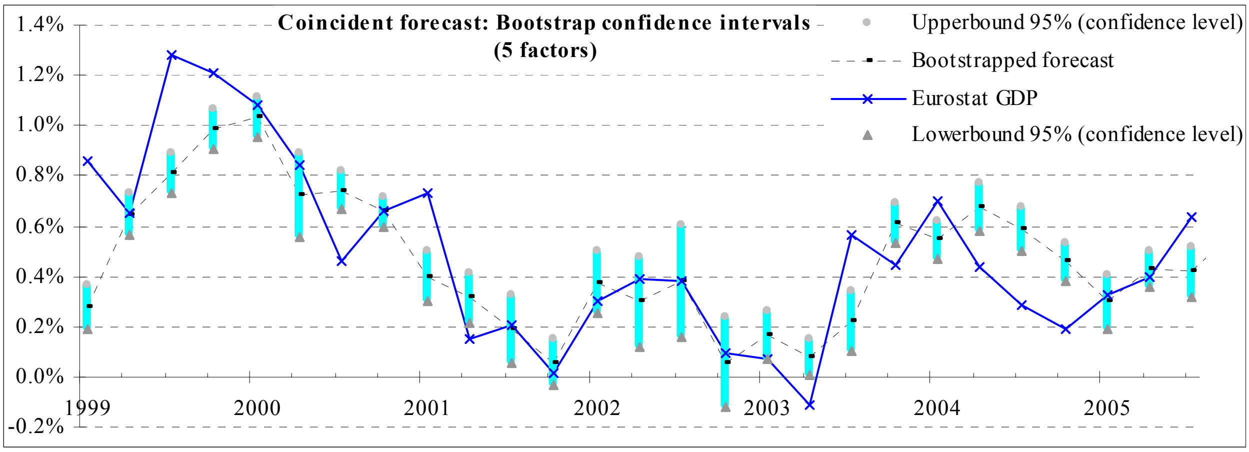


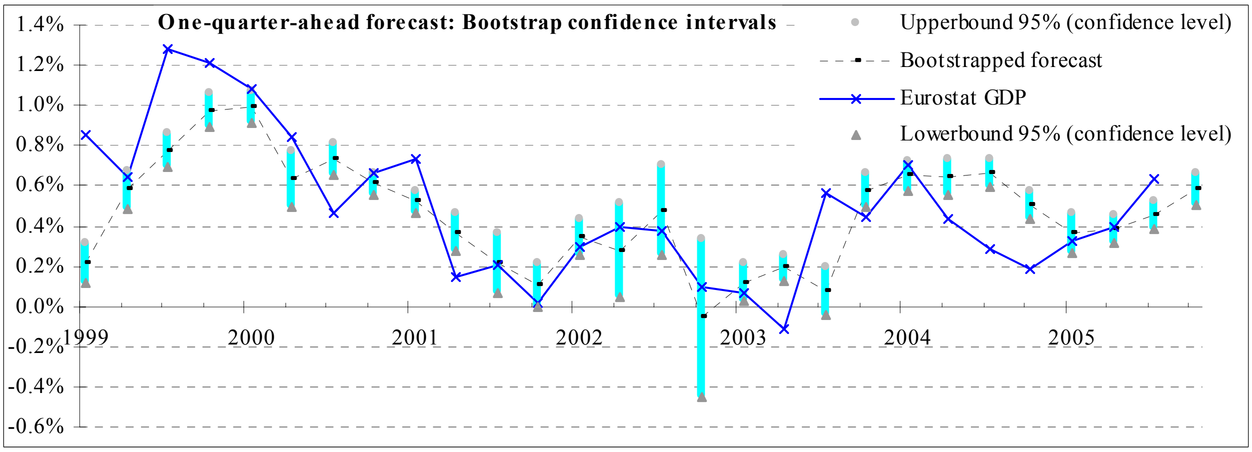


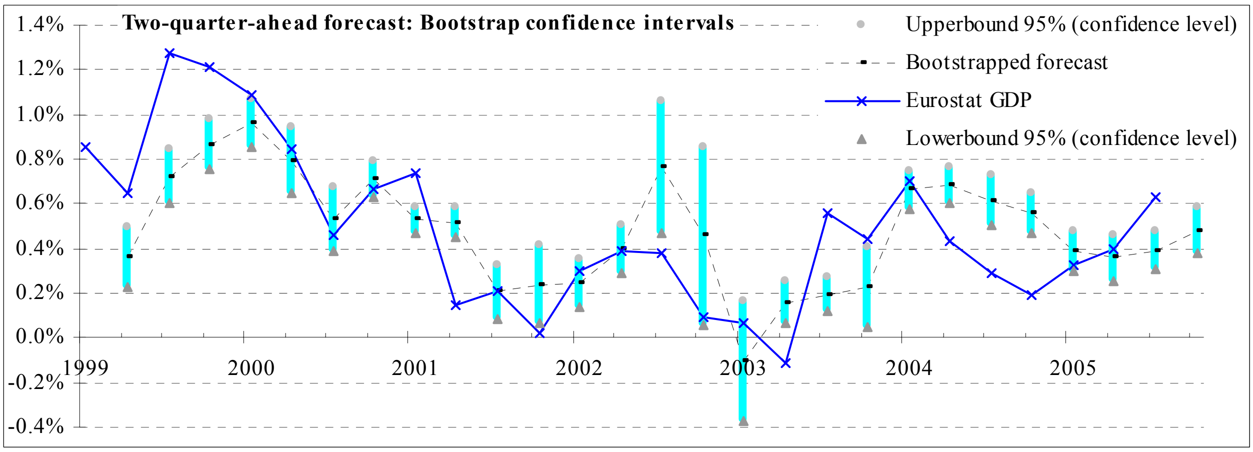


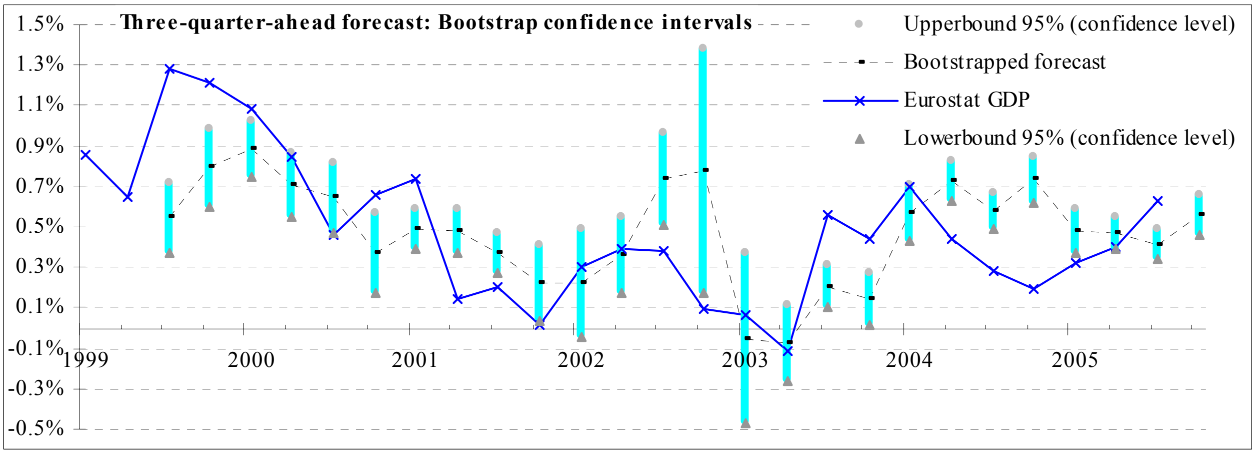


◊ *Forecasts confidence intervals with five factors*

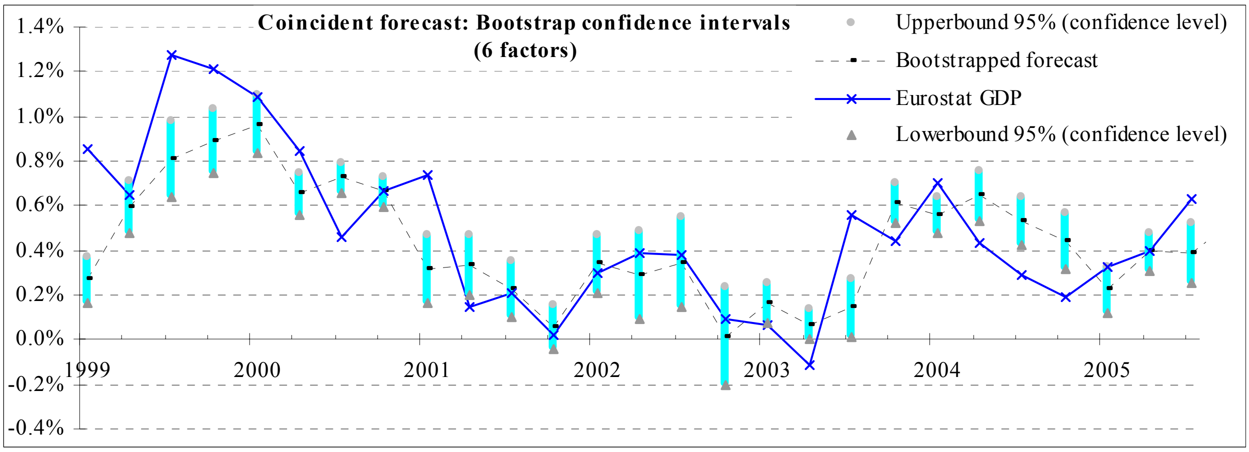


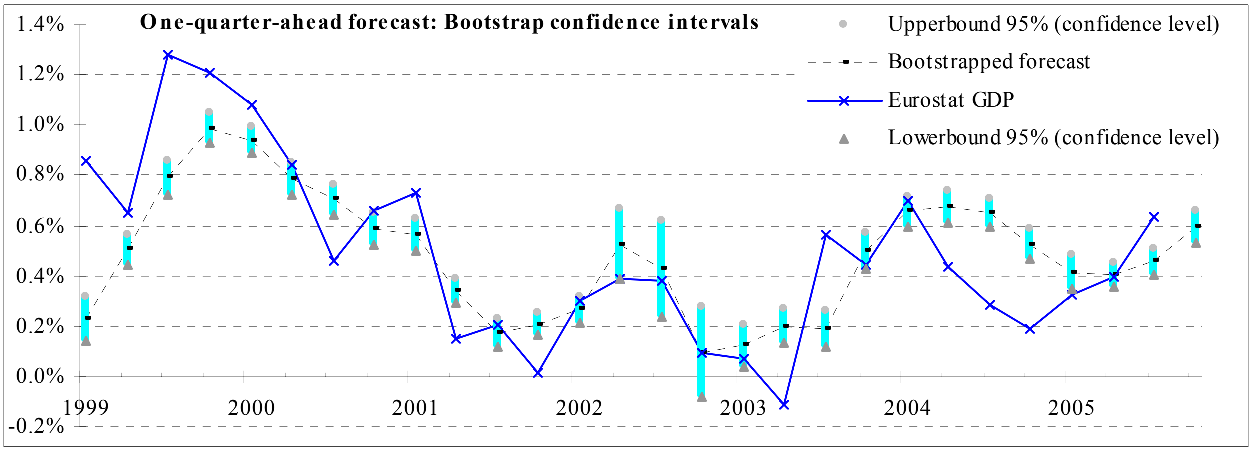


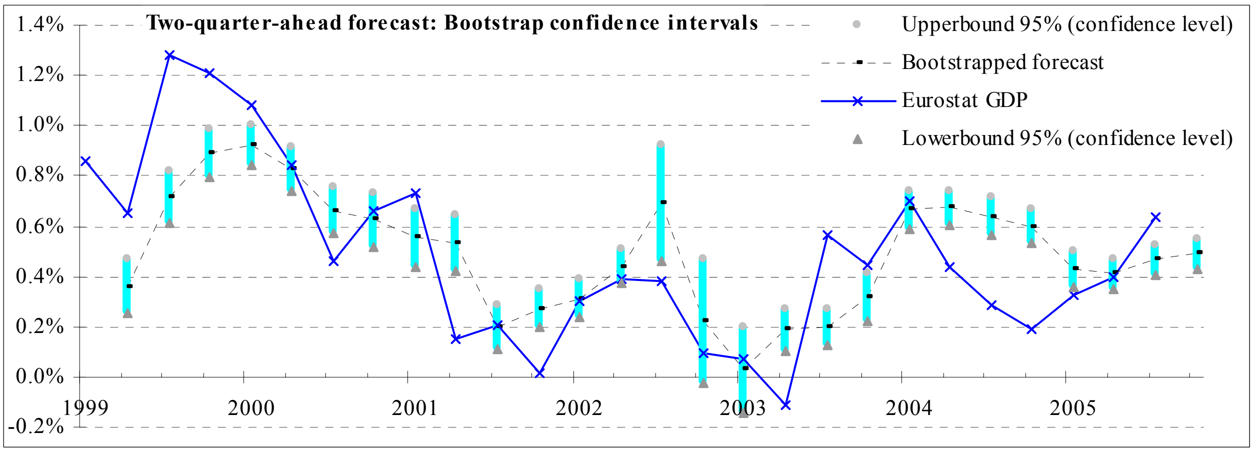


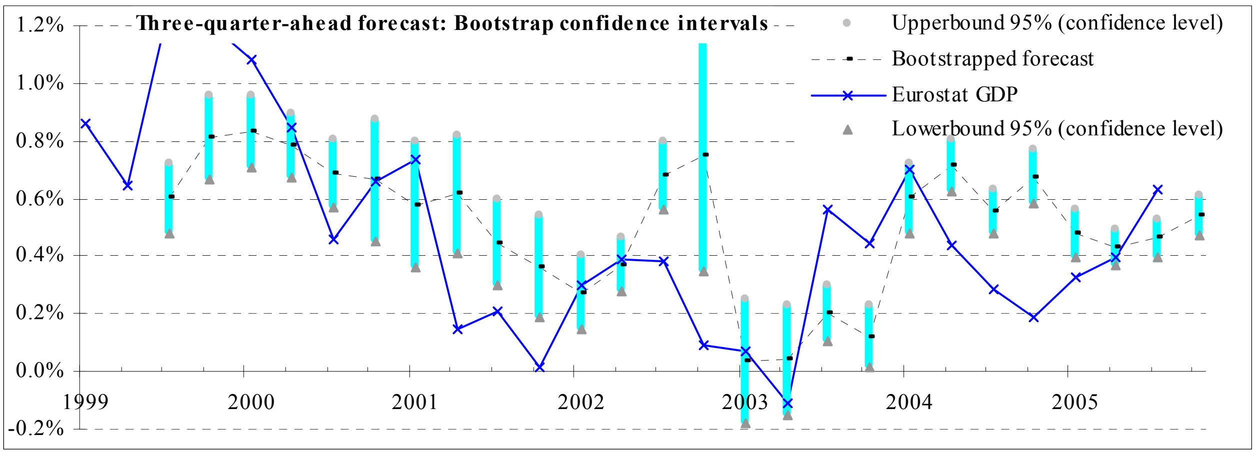


◊ *Forecasts confidence intervals with six factors*

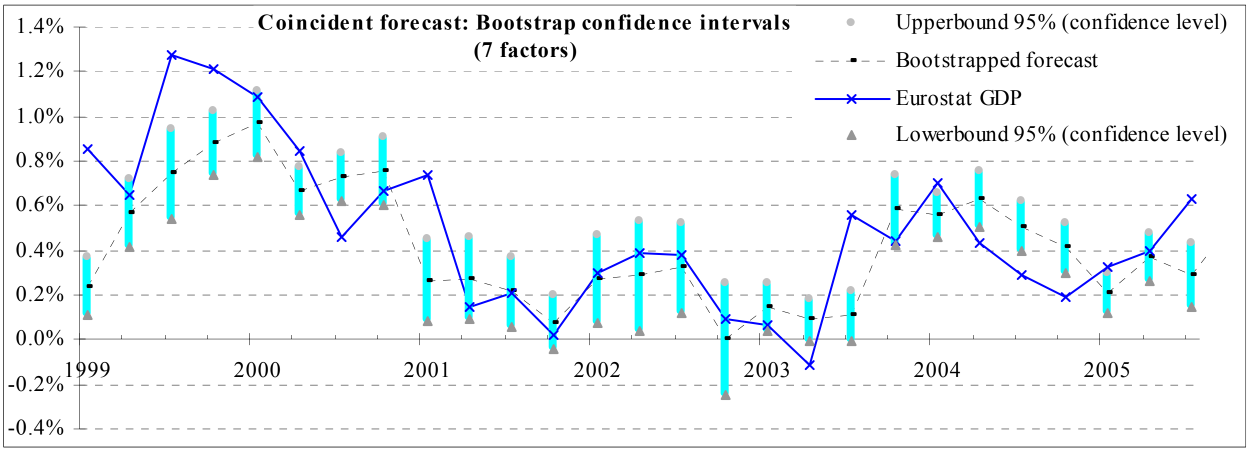


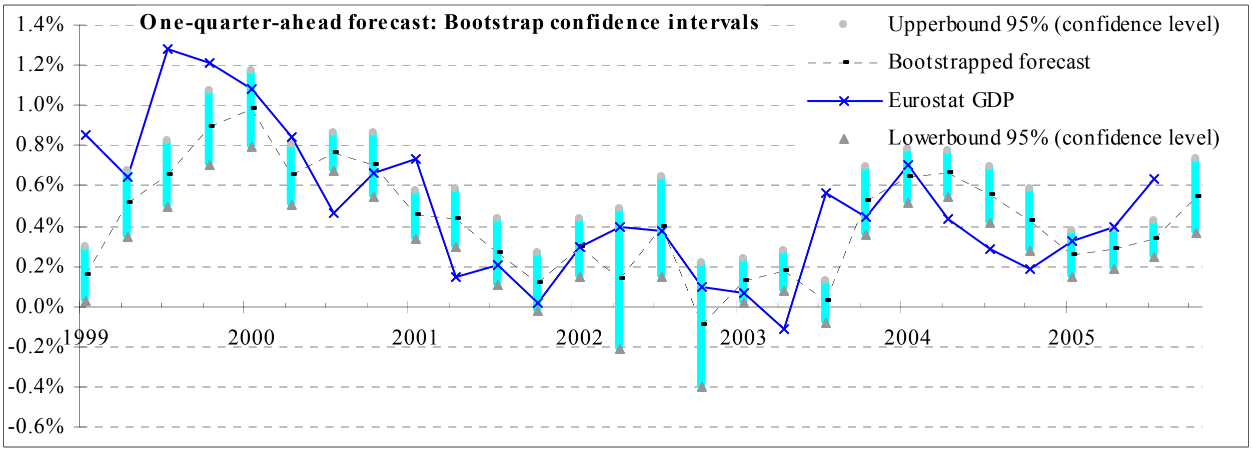


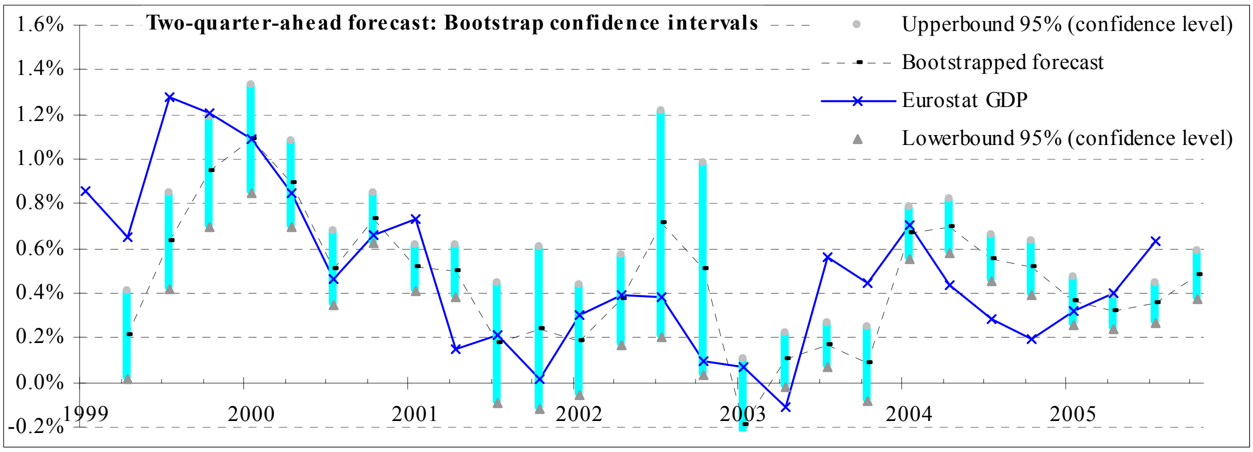


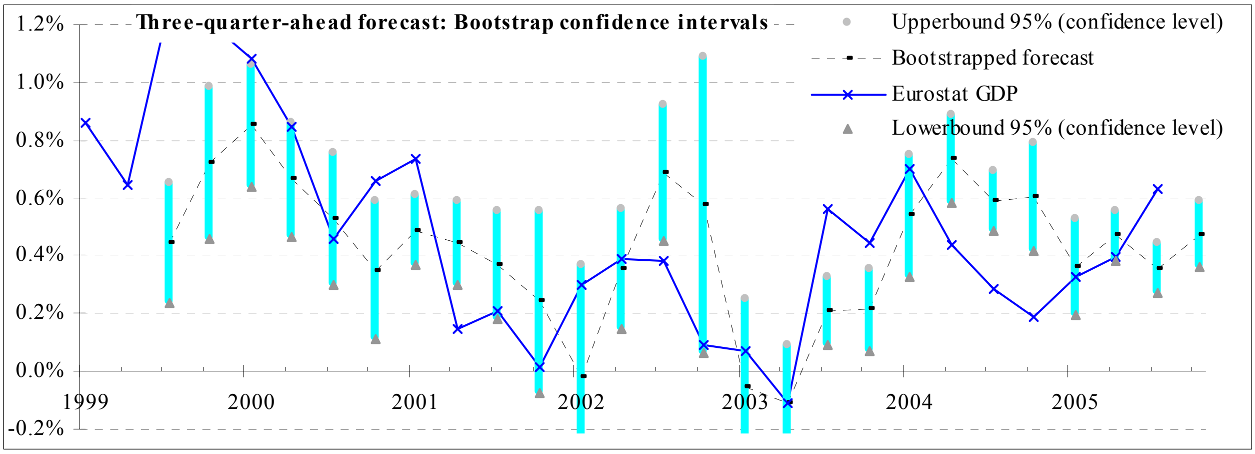


◊ *Forecasts confidence intervals with seven factors*









ANNEX 4: INFORMATION CRITERIA (IC)

4.1. The three panel information criteria of Bai and Ng (2002)

◊ *The criteria*

Bai and Ng (2002) tackle the issue of the estimation of the number of factors as a problem of model selection, where each model allows for a different number of latent factors. They suggest using three information criteria (BNIC) based on the residuals of the time-series regressions of predictors on a given set of k factors, corrected by a penalty term.

$$IC_{p1}(k) = \ln\left(\underset{\Lambda=(\lambda^1,\ldots\lambda^k)}{Min} \left\{ \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{it} - \lambda_i^{k'} \hat{F}_t^k)^2 \right\} \right) + k\left(\frac{N+T}{NT} \right) \ln\left(\frac{NT}{N+T} \right)$$

$$IC_{p2}(k) = \ln\left(\underset{\Lambda=(\lambda^1,\ldots\lambda^k)}{Min} \left\{ \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{it} - \lambda_i^{k'} \hat{F}_t^k)^2 \right\} \right) + k\left(\frac{N+T}{NT} \right) \ln\left(Min\{N,T\} \right)$$

$$IC_{p3}(k) = \ln\left(\underset{\Lambda=(\lambda^1,\ldots\lambda^k)}{Min} \left\{ \frac{1}{NT} \sum_{i=1}^{N} \sum_{t=1}^{T} (X_{it} - \lambda_i^{k'} \hat{F}_t^k)^2 \right\} \right) + k\left(\frac{\ln\left(Min\{N,T\} \right)}{Min\{N,T\}} \right)$$

The BNIC computed in the same pseudo-real time framework as that of forecasts always prescribes to select only the first eigenvector irrespective of out-of-sample quarters or forecast horizons. Given anecdotal evidence on eigenvalues distribution or out-of-sample forecast accuracy statistics, the penalty term in the BNIC seems severely over-calibrated for real data applications.

◊ *Results: coincident forecasts*

Coincident forecast		Bai and Ng IC_{p1}							Bai and Ng IC_{p2}							Bai and Ng IC_{p3}						
factors		1	2	3	4	5	6	min	1	2	3	4	5	6	min	1	2	3	4	5	6	min
scores quarter by quarter	99Q1	15.47	15.55	15.65	15.70	15.72	15.80	**1**	15.47	15.55	15.65	15.70	15.72	15.80	**1**	15.47	15.55	15.65	15.69	15.72	15.80	**1**
	99Q2	15.45	15.52	15.62	15.66	15.71	15.80	**1**	15.45	15.52	15.63	15.66	15.71	15.80	**1**	15.45	15.52	15.62	15.65	15.71	15.79	**1**
	99Q3	15.42	15.51	15.60	15.64	15.70	15.77	**1**	15.42	15.51	15.60	15.64	15.70	15.78	**1**	15.42	15.51	15.60	15.64	15.69	15.77	**1**
	99Q4	15.43	15.52	15.60	15.65	15.70	15.76	**1**	15.43	15.52	15.60	15.65	15.70	15.76	**1**	15.43	15.52	15.60	15.64	15.70	15.76	**1**
	00Q1	15.46	15.56	15.65	15.69	15.75	15.82	**1**	15.46	15.56	15.65	15.69	15.75	15.83	**1**	15.46	15.55	15.64	15.69	15.75	15.82	**1**
	00Q2	15.49	15.58	15.68	15.73	15.78	15.87	**1**	15.49	15.58	15.68	15.73	15.78	15.87	**1**	15.49	15.58	15.68	15.73	15.78	15.87	**1**
	00Q3	15.50	15.59	15.69	15.76	15.82	15.86	**1**	15.50	15.59	15.69	15.76	15.82	15.86	**1**	15.50	15.59	15.69	15.76	15.82	15.86	**1**
	00Q4	15.55	15.64	15.74	15.84	15.87	15.88	**1**	15.55	15.64	15.74	15.84	15.88	15.88	**1**	15.55	15.64	15.74	15.84	15.87	15.88	**1**
	01Q1	15.59	15.68	15.79	15.88	15.87	15.93	**1**	15.59	15.68	15.79	15.88	15.87	15.93	**1**	15.59	15.68	15.78	15.88	15.87	15.93	**1**
	01Q2	15.65	15.74	15.85	15.95	15.93	15.94	**1**	15.65	15.74	15.85	15.95	15.93	15.94	**1**	15.65	15.74	15.84	15.95	15.93	15.93	**1**
	01Q3	15.71	15.81	15.91	16.00	16.01	16.04	**1**	15.71	15.81	15.91	16.00	16.01	16.05	**1**	15.71	15.81	15.91	16.00	16.01	16.04	**1**
	01Q4	15.75	15.85	15.96	16.05	16.06	16.06	**1**	15.75	15.85	15.96	16.05	16.06	16.07	**1**	15.75	15.85	15.96	16.05	16.06	16.06	**1**
	02Q1	15.78	15.88	15.99	16.04	16.11	16.13	**1**	15.78	15.88	15.99	16.04	16.11	16.13	**1**	15.78	15.88	15.98	16.04	16.10	16.12	**1**
	02Q2	15.80	15.90	16.01	16.06	16.12	16.18	**1**	15.80	15.90	16.01	16.06	16.12	16.18	**1**	15.80	15.90	16.01	16.06	16.12	16.17	**1**
	02Q3	16.22	16.32	16.39	16.44	16.54	16.60	**1**	16.22	16.32	16.39	16.44	16.54	16.61	**1**	16.22	16.32	16.38	16.44	16.54	16.60	**1**
	02Q4	16.48	16.59	16.69	16.74	16.76	16.86	**1**	16.48	16.59	16.69	16.75	16.76	16.86	**1**	16.48	16.59	16.69	16.74	16.76	16.86	**1**
	03Q1	16.68	16.78	16.88	16.96	16.99	17.06	**1**	16.68	16.78	16.88	16.96	16.99	17.07	**1**	16.67	16.78	16.88	16.96	16.99	17.06	**1**
	03Q2	16.83	16.94	17.05	17.12	17.16	17.24	**1**	16.83	16.94	17.05	17.12	17.16	17.24	**1**	16.83	16.94	17.05	17.12	17.16	17.24	**1**
	03Q3	16.82	16.93	17.04	17.10	17.14	17.23	**1**	16.82	16.93	17.04	17.10	17.14	17.23	**1**	16.82	16.93	17.04	17.10	17.13	17.23	**1**
	03Q4	16.83	16.94	17.05	17.12	17.15	17.21	**1**	16.83	16.94	17.05	17.12	17.15	17.21	**1**	16.83	16.94	17.05	17.12	17.15	17.21	**1**
	04Q1	16.87	16.97	17.08	17.14	17.18	17.23	**1**	16.87	16.97	17.08	17.14	17.18	17.23	**1**	16.87	16.97	17.08	17.14	17.17	17.23	**1**
	04Q2	16.90	17.00	17.11	17.17	17.22	17.26	**1**	16.90	17.00	17.11	17.17	17.22	17.26	**1**	16.90	17.00	17.11	17.17	17.22	17.26	**1**
	04Q3	16.93	17.04	17.14	17.21	17.26	17.31	**1**	16.93	17.04	17.15	17.21	17.26	17.31	**1**	16.93	17.04	17.14	17.21	17.26	17.31	**1**
	04Q4	16.96	17.06	17.17	17.24	17.29	17.36	**1**	16.96	17.06	17.17	17.24	17.29	17.36	**1**	16.96	17.06	17.17	17.24	17.29	17.36	**1**
	05Q1	16.96	17.07	17.18	17.24	17.29	17.38	**1**	16.96	17.07	17.18	17.24	17.29	17.38	**1**	16.96	17.07	17.17	17.24	17.29	17.38	**1**
	05Q2	16.95	17.05	17.16	17.23	17.29	17.37	**1**	16.95	17.05	17.16	17.24	17.29	17.37	**1**	16.95	17.05	17.16	17.23	17.29	17.37	**1**
	05Q3	16.94	17.05	17.16	17.22	17.29	17.37	**1**	16.94	17.05	17.16	17.22	17.29	17.37	**1**	16.94	17.05	17.16	17.22	17.28	17.37	**1**
average score		**16.16**	**16.26**	**16.36**	**16.43**	**16.47**	**16.53**	**1** **100%**	**16.16**	**16.26**	**16.36**	**16.43**	**16.47**	**16.53**	**1** **100%**	**16.16**	**16.26**	**16.36**	**16.43**	**16.47**	**16.53**	**1** **100%**

◊ *Results: one-quarter-ahead forecasts*

1Q ahead forecast		Bai and Ng IC_{p1}							Bai and Ng IC_{p2}							Bai and Ng IC_{p3}						
factors		1	2	3	4	5	6	min	1	2	3	4	5	6	min	1	2	3	4	5	6	min
	99Q1	15.27	15.38	15.47	15.51	15.50	15.59	**1**	15.27	15.38	15.47	15.52	15.50	15.59	**1**	15.27	15.37	15.47	15.51	15.50	15.59	**1**
	99Q2	15.32	15.41	15.51	15.54	15.59	15.65	**1**	15.32	15.41	15.51	15.54	15.59	15.66	**1**	15.32	15.41	15.51	15.53	15.58	15.65	**1**
	99Q3	15.27	15.35	15.45	15.48	15.53	15.63	**1**	15.27	15.35	15.45	15.48	15.53	15.63	**1**	15.27	15.35	15.45	15.48	15.52	15.62	**1**
	99Q4	15.30	15.39	15.47	15.51	15.56	15.65	**1**	15.30	15.39	15.47	15.51	15.56	15.65	**1**	15.30	15.38	15.47	15.51	15.56	15.64	**1**
	00Q1	15.31	15.40	15.48	15.51	15.58	15.58	**1**	15.31	15.40	15.48	15.51	15.58	15.58	**1**	15.30	15.40	15.48	15.51	15.57	15.58	**1**
	00Q2	15.34	15.43	15.53	15.58	15.62	15.70	**1**	15.34	15.43	15.53	15.58	15.62	15.70	**1**	15.34	15.43	15.53	15.57	15.61	15.69	**1**
	00Q3	15.36	15.46	15.56	15.60	15.67	15.73	**1**	15.36	15.46	15.56	15.60	15.67	15.73	**1**	15.36	15.45	15.55	15.60	15.66	15.73	**1**
	00Q4	15.39	15.49	15.59	15.69	15.78	15.71	**1**	15.39	15.49	15.59	15.69	15.78	15.71	**1**	15.39	15.49	15.59	15.69	15.77	15.71	**1**
	01Q1	15.42	15.51	15.61	15.71	15.67	15.74	**1**	15.42	15.51	15.62	15.71	15.67	15.74	**1**	15.42	15.51	15.61	15.71	15.67	15.74	**1**
	01Q2	15.48	15.57	15.67	15.78	15.74	15.82	**1**	15.48	15.57	15.68	15.78	15.74	15.82	**1**	15.48	15.57	15.67	15.78	15.73	15.82	**1**
	01Q3	15.56	15.65	15.76	15.86	15.85	15.86	**1**	15.56	15.65	15.76	15.86	15.85	15.86	**1**	15.56	15.65	15.75	15.86	15.84	15.85	**1**
	01Q4	15.61	15.71	15.82	15.91	15.91	15.94	**1**	15.61	15.71	15.82	15.91	15.91	15.94	**1**	15.61	15.71	15.82	15.91	15.90	15.93	**1**
scores	02Q1	15.64	15.74	15.85	15.92	15.95	15.99	**1**	15.64	15.74	15.85	15.92	15.95	15.99	**1**	15.64	15.74	15.85	15.92	15.95	15.98	**1**
quarter	02Q2	15.65	15.76	15.86	15.90	15.95	16.03	**1**	15.65	15.76	15.86	15.90	15.95	16.03	**1**	15.65	15.76	15.86	15.90	15.95	16.03	**1**
by	02Q3	15.68	15.78	15.86	15.93	15.98	16.05	**1**	15.68	15.78	15.86	15.93	15.99	16.05	**1**	15.68	15.78	15.86	15.92	15.98	16.04	**1**
quarter	02Q4	16.18	16.29	16.38	16.42	16.53	16.55	**1**	16.19	16.29	16.38	16.43	16.53	16.55	**1**	16.18	16.29	16.38	16.42	16.53	16.55	**1**
	03Q1	16.49	16.59	16.69	16.75	16.81	16.83	**1**	16.49	16.59	16.69	16.75	16.81	16.84	**1**	16.49	16.59	16.69	16.75	16.81	16.83	**1**
	03Q2	16.69	16.79	16.90	16.96	17.03	17.11	**1**	16.69	16.80	16.90	16.96	17.04	17.11	**1**	16.69	16.79	16.90	16.95	17.03	17.10	**1**
	03Q3	16.69	16.79	16.90	16.96	17.05	17.12	**1**	16.69	16.79	16.90	16.96	17.05	17.12	**1**	16.68	16.79	16.90	16.96	17.05	17.12	**1**
	03Q4	16.70	16.80	16.91	16.98	17.05	17.09	**1**	16.70	16.80	16.91	16.98	17.05	17.09	**1**	16.69	16.80	16.91	16.98	17.05	17.09	**1**
	04Q1	16.71	16.81	16.92	16.99	17.07	17.07	**1**	16.71	16.81	16.92	16.99	17.07	17.08	**1**	16.71	16.81	16.92	16.99	17.06	17.07	**1**
	04Q2	16.74	16.84	16.94	17.02	17.10	17.11	**1**	16.74	16.84	16.94	17.02	17.10	17.11	**1**	16.74	16.84	16.94	17.02	17.10	17.10	**1**
	04Q3	16.78	16.88	16.99	17.06	17.15	17.16	**1**	16.78	16.88	16.99	17.06	17.15	17.16	**1**	16.78	16.88	16.99	17.06	17.14	17.16	**1**
	04Q4	16.82	16.93	17.04	17.10	17.19	17.21	**1**	16.82	16.93	17.04	17.10	17.19	17.21	**1**	16.82	16.93	17.03	17.10	17.18	17.21	**1**
	05Q1	16.83	16.93	17.04	17.11	17.19	17.22	**1**	16.83	16.93	17.04	17.11	17.19	17.22	**1**	16.83	16.93	17.04	17.11	17.19	17.21	**1**
	05Q2	16.82	16.92	17.03	17.10	17.18	17.21	**1**	16.82	16.92	17.03	17.10	17.18	17.21	**1**	16.82	16.92	17.03	17.10	17.18	17.21	**1**
	05Q3	16.80	16.91	17.02	17.08	17.17	17.21	**1**	16.80	16.91	17.02	17.08	17.17	17.21	**1**	16.80	16.91	17.02	17.08	17.17	17.21	**1**
	05Q4	16.80	16.90	17.01	17.08	17.17	17.23	**1**	16.80	16.90	17.01	17.08	17.17	17.23	**1**	16.80	16.90	17.01	17.08	17.17	17.23	**1**
average score		**15.99**	**16.09**	**16.19**	**16.26**	**16.31**	**16.35**	**1** **100%**	**15.99**	**16.09**	**16.20**	**16.26**	**16.31**	**16.35**	**1** **100%**	**15.99**	**16.09**	**16.19**	**16.26**	**16.31**	**16.35**	**1** **100%**

◊ *Results: two-quarter-ahead forecasts*

2Q ahead forecast		Bai and Ng IC_{p1}							Bai and Ng IC_{p2}							Bai and Ng IC_{p3}						
factors		1	2	3	4	5	6	min	1	2	3	4	5	6	min	1	2	3	4	5	6	min
	99Q2	15.22	15.32	15.42	15.46	15.44	15.53	**1**	15.22	15.32	15.42	15.46	15.44	15.53	**1**	15.22	15.32	15.42	15.46	15.43	15.52	**1**
	99Q3	15.17	15.27	15.37	15.41	15.41	15.52	**1**	15.17	15.27	15.37	15.41	15.41	15.52	**1**	15.17	15.27	15.36	15.40	15.41	15.51	**1**
	99Q4	15.18	15.26	15.36	15.40	15.40	15.47	**1**	15.18	15.27	15.36	15.40	15.40	15.48	**1**	15.18	15.26	15.36	15.40	15.39	15.47	**1**
	00Q1	15.19	15.27	15.37	15.42	15.40	15.50	**1**	15.19	15.27	15.37	15.42	15.40	15.51	**1**	15.19	15.27	15.36	15.42	15.39	15.50	**1**
	00Q2	15.18	15.28	15.37	15.43	15.41	15.47	**1**	15.18	15.28	15.37	15.43	15.42	15.47	**1**	15.18	15.27	15.37	15.42	15.41	15.47	**1**
	00Q3	15.19	15.29	15.38	15.44	15.41	15.43	**1**	15.19	15.29	15.39	15.44	15.41	15.43	**1**	15.19	15.28	15.38	15.44	15.41	15.42	**1**
	00Q4	15.26	15.36	15.46	15.54	15.54	15.53	**1**	15.26	15.36	15.46	15.54	15.54	15.53	**1**	15.26	15.36	15.45	15.54	15.53	15.52	**1**
	01Q1	15.31	15.42	15.52	15.61	15.70	15.59	**1**	15.31	15.42	15.52	15.61	15.70	15.59	**1**	15.31	15.42	15.52	15.61	15.70	15.58	**1**
	01Q2	15.30	15.40	15.50	15.59	15.47	15.55	**1**	15.30	15.40	15.50	15.60	15.48	15.55	**1**	15.30	15.40	15.49	15.59	15.47	15.55	**1**
	01Q3	15.35	15.45	15.55	15.64	15.58	15.65	**1**	15.36	15.46	15.55	15.65	15.58	15.65	**1**	15.35	15.45	15.54	15.64	15.58	15.65	**1**
	01Q4	15.46	15.56	15.66	15.75	15.72	15.78	**1**	15.46	15.56	15.66	15.76	15.72	15.78	**1**	15.45	15.56	15.66	15.75	15.71	15.77	**1**
	02Q1	15.51	15.61	15.71	15.79	15.78	15.84	**1**	15.51	15.61	15.72	15.80	15.78	15.84	**1**	15.51	15.61	15.71	15.79	15.77	15.83	**1**
scores	02Q2	15.55	15.65	15.75	15.80	15.83	15.90	**1**	15.55	15.65	15.76	15.80	15.83	15.91	**1**	15.55	15.65	15.75	15.79	15.82	15.90	**1**
quarter	02Q3	15.57	15.68	15.76	15.80	15.86	15.94	**1**	15.57	15.68	15.76	15.80	15.86	15.94	**1**	15.57	15.67	15.76	15.80	15.85	15.93	**1**
by	02Q4	15.59	15.69	15.79	15.79	15.87	15.95	**1**	15.59	15.69	15.79	15.79	15.87	15.95	**1**	15.59	15.69	15.78	15.78	15.87	15.94	**1**
quarter	03Q1	16.23	16.31	16.39	16.44	16.50	16.57	**1**	16.23	16.31	16.39	16.44	16.50	16.57	**1**	16.23	16.31	16.39	16.43	16.50	16.56	**1**
	03Q2	16.60	16.69	16.79	16.78	16.89	16.91	**1**	16.60	16.69	16.79	16.78	16.89	16.91	**1**	16.59	16.69	16.79	16.78	16.89	16.90	**1**
	03Q3	16.60	16.69	16.79	16.85	16.93	16.93	**1**	16.60	16.69	16.79	16.85	16.93	16.93	**1**	16.60	16.69	16.79	16.84	16.93	16.92	**1**
	03Q4	16.60	16.70	16.80	16.86	16.93	16.95	**1**	16.60	16.70	16.80	16.86	16.93	16.96	**1**	16.60	16.69	16.80	16.86	16.92	16.95	**1**
	04Q1	16.62	16.72	16.82	16.88	16.90	16.93	**1**	16.62	16.72	16.82	16.88	16.90	16.93	**1**	16.62	16.72	16.82	16.88	16.89	16.92	**1**
	04Q2	16.64	16.73	16.83	16.91	16.91	16.98	**1**	16.64	16.73	16.83	16.92	16.91	16.98	**1**	16.64	16.73	16.83	16.91	16.90	16.98	**1**
	04Q3	16.64	16.73	16.83	16.93	16.91	17.00	**1**	16.64	16.73	16.84	16.93	16.92	17.00	**1**	16.64	16.73	16.83	16.92	16.91	17.00	**1**
	04Q4	16.69	16.78	16.88	16.97	16.97	17.05	**1**	16.69	16.78	16.88	16.97	16.97	17.05	**1**	16.69	16.78	16.88	16.97	16.97	17.05	**1**
	05Q1	16.75	16.84	16.94	17.03	17.05	17.12	**1**	16.75	16.84	16.95	17.03	17.05	17.12	**1**	16.75	16.84	16.94	17.03	17.05	17.12	**1**
	05Q2	16.77	16.86	16.96	17.04	17.10	17.14	**1**	16.77	16.86	16.96	17.05	17.10	17.14	**1**	16.76	16.85	16.96	17.04	17.09	17.14	**1**
	05Q3	16.75	16.84	16.95	17.02	17.09	17.13	**1**	16.75	16.84	16.95	17.02	17.09	17.13	**1**	16.75	16.84	16.94	17.02	17.09	17.13	**1**
	05Q4	16.74	16.82	16.93	17.01	17.10	17.12	**1**	16.74	16.82	16.93	17.01	17.10	17.12	**1**	16.74	16.82	16.93	17.00	17.10	17.11	**1**
	06Q1	16.73	16.81	16.92	17.01	17.11	17.14	**1**	16.73	16.82	16.92	17.01	17.11	17.14	**1**	16.73	16.81	16.92	17.01	17.11	17.13	**1**
average score		**15.94**	**16.04**	**16.14**	**16.20**	**16.22**	**16.27**	**1** **100%**	**15.94**	**16.04**	**16.14**	**16.20**	**16.22**	**16.27**	**1** **100%**	**15.94**	**16.03**	**16.13**	**16.20**	**16.22**	**16.27**	**1** **100%**

◊ *Results: three-quarter-ahead forecasts*

3Q ahead forecast		Bai and Ng IC_{p1}							Bai and Ng IC_{p2}							Bai and Ng IC_{p3}						
factors		1	2	3	4	5	6	min	1	2	3	4	5	6	min	1	2	3	4	5	6	min
	99Q3	15.27	15.37	15.47	15.51	15.50	15.59	**1**	15.22	15.32	15.42	15.46	15.43	15.52	**1**	14.93	15.03	15.13	15.21	15.21	15.31	**1**
	99Q4	15.32	15.41	15.51	15.53	15.58	15.65	**1**	15.17	15.27	15.36	15.40	15.41	15.51	**1**	14.98	15.08	15.17	15.25	15.22	15.33	**1**
	00Q1	15.27	15.35	15.45	15.48	15.52	15.62	**1**	15.18	15.26	15.36	15.40	15.39	15.47	**1**	15.00	15.11	15.20	15.28	15.34	15.30	**1**
	00Q2	15.30	15.38	15.47	15.51	15.56	15.64	**1**	15.19	15.27	15.36	15.42	15.39	15.50	**1**	15.02	15.11	15.20	15.28	15.35	15.31	**1**
	00Q3	15.30	15.40	15.48	15.51	15.57	15.58	**1**	15.18	15.27	15.37	15.42	15.41	15.47	**1**	14.92	15.01	15.10	15.20	15.19	15.22	**1**
	00Q4	15.34	15.43	15.53	15.57	15.61	15.69	**1**	15.19	15.28	15.38	15.44	15.41	15.42	**1**	15.01	15.11	15.21	15.28	15.37	15.14	**1**
	01Q1	15.36	15.45	15.55	15.60	15.66	15.73	**1**	15.26	15.36	15.45	15.54	15.53	15.52	**1**	15.00	15.11	15.19	15.30	15.30	15.16	**1**
	01Q2	15.39	15.49	15.59	15.69	15.77	15.71	**1**	15.31	15.42	15.52	15.61	15.70	15.58	**1**	15.14	15.25	15.34	15.44	15.45	15.34	**1**
	01Q3	15.42	15.51	15.61	15.71	15.67	15.74	**1**	15.30	15.40	15.49	15.59	15.47	15.55	**1**	15.13	15.24	15.29	15.40	15.26	15.31	**1**
	01Q4	15.48	15.57	15.67	15.78	15.73	15.82	**1**	15.35	15.45	15.54	15.64	15.58	15.65	**1**	15.11	15.21	15.30	15.40	15.31	15.36	**1**
	02Q1	15.56	15.65	15.75	15.86	15.84	15.85	**1**	15.45	15.56	15.66	15.75	15.71	15.77	**1**	15.25	15.35	15.46	15.53	15.51	15.54	**1**
	02Q2	15.61	15.71	15.82	15.91	15.90	15.93	**1**	15.51	15.61	15.71	15.79	15.77	15.83	**1**	15.30	15.39	15.48	15.55	15.59	15.66	**1**
scores	02Q3	15.64	15.74	15.85	15.92	15.95	15.98	**1**	15.55	15.65	15.75	15.79	15.82	15.90	**1**	15.34	15.44	15.53	15.62	15.64	15.71	**1**
quarter	02Q4	15.65	15.76	15.86	15.90	15.95	16.03	**1**	15.57	15.67	15.76	15.80	15.85	15.93	**1**	15.40	15.50	15.56	15.66	15.68	15.76	**1**
by	03Q1	15.68	15.78	15.86	15.92	15.98	16.04	**1**	15.59	15.69	15.78	15.78	15.87	15.94	**1**	15.41	15.51	15.56	15.64	15.72	15.72	**1**
quarter	03Q2	16.18	16.29	16.38	16.42	16.53	16.55	**1**	16.23	16.31	16.39	16.43	16.50	16.56	**1**	16.26	16.37	16.42	16.53	16.56	16.53	**1**
	03Q3	16.49	16.59	16.69	16.75	16.81	16.83	**1**	16.59	16.69	16.79	16.78	16.89	16.90	**1**	16.37	16.42	16.52	16.61	16.56	16.66	**1**
	03Q4	16.69	16.79	16.90	16.95	17.03	17.10	**1**	16.60	16.69	16.79	16.84	16.93	16.92	**1**	16.43	16.51	16.60	16.65	16.70	16.74	**1**
	04Q1	16.68	16.79	16.90	16.96	17.05	17.12	**1**	16.60	16.69	16.80	16.86	16.92	16.95	**1**	16.43	16.53	16.61	16.62	16.68	16.74	**1**
	04Q2	16.69	16.80	16.91	16.98	17.05	17.09	**1**	16.62	16.72	16.82	16.88	16.89	16.92	**1**	16.42	16.53	16.60	16.58	16.66	16.74	**1**
	04Q3	16.71	16.81	16.92	16.99	17.06	17.07	**1**	16.64	16.73	16.83	16.91	16.90	16.98	**1**	16.47	16.57	16.64	16.66	16.74	16.81	**1**
	04Q4	16.74	16.84	16.94	17.02	17.10	17.10	**1**	16.64	16.73	16.83	16.92	16.91	17.00	**1**	16.48	16.58	16.66	16.71	16.76	16.84	**1**
	05Q1	16.78	16.88	16.99	17.06	17.14	17.16	**1**	16.69	16.78	16.88	16.97	16.97	17.05	**1**	16.48	16.57	16.65	16.69	16.74	16.84	**1**
	05Q2	16.82	16.93	17.03	17.10	17.18	17.21	**1**	16.75	16.84	16.94	17.03	17.05	17.12	**1**	16.56	16.66	16.75	16.81	16.83	16.92	**1**
	05Q3	16.83	16.93	17.04	17.11	17.19	17.21	**1**	16.76	16.85	16.96	17.04	17.09	17.14	**1**	16.60	16.69	16.78	16.87	16.89	17.00	**1**
	05Q4	16.82	16.92	17.03	17.10	17.18	17.21	**1**	16.75	16.84	16.94	17.02	17.09	17.13	**1**	16.59	16.69	16.77	16.86	16.88	16.95	**1**
	06Q1	16.80	16.91	17.02	17.08	17.17	17.21	**1**	16.74	16.82	16.93	17.00	17.10	17.11	**1**	16.57	16.66	16.74	16.84	16.87	16.96	**1**
	06Q2	16.80	16.90	17.01	17.08	17.17	17.23	**1**	16.73	16.81	16.92	17.01	17.11	17.13	**1**	16.56	16.65	16.74	16.84	16.88	16.96	**1**
average score		**16.02**	**16.12**	**16.22**	**16.29**	**16.34**	**16.38**	**1** **100%**	**15.94**	**16.03**	**16.13**	**16.20**	**16.22**	**16.27**	**1** **100%**	**15.76**	**15.85**	**15.94**	**16.01**	**16.03**	**16.07**	**1** **100%**

4.2. The Bayes Information Criterion (BIC)

◊ *The adapted BIC for factor selection*

Given that the criteria seem well suited to series with a high signal to noise ratio, a natural solution for the SLID model[85] is to use a classic BIC based on the fit of the endogenous series (GDP) as measured by the SSR of the OLS regression of GDP on subsets of factors.

$$IC_{GDP3}(k) = \ln\left(\underset{\Lambda=(\lambda^1,\ldots\lambda^k)}{Min} \left\{ \frac{1}{NT} \sum_{t=1}^{T} (GDP_t - \lambda_i^{k'} \hat{F}_t^k)^2 \right\} \right) + k \left(\frac{\ln(Min\{N,T\})}{Min\{N,T\}} \right)$$

As in the case of the BNIC, the time spans used for the regressions end at the last observation available for GDP, meaning that they do neither include GDP forecasts (coincident or at more remote horizons) nor factors based on the predictors' observations corresponding to the forecasts. This is motivated by the fact that forecasts rely on the choice of a specific subset of factors.

Three factors are suggested by the adapted BIC based on the average score across quarters and for about 2/3 of the out-of-sample quarters (except in the case of three-quarter-ahead forecasts where results quarter by quarter seem more noisy). Results are far more consistent with the monitoring of the decay of the eigenvalues and out-of-sample empirical results are those derived from using the BNIC criteria.

85 Recall that forecasts are computed with a projection of GDP on the coincident factors and not on any lagged terms.

◊ *Results: coincident forecasts*

factors	Coincident forecast: BIC IC_{GDP3} 1	2	3	4	5	6	min
99Q1	-19.20	-19.17	-19.28	-19.24	-19.18	-19.06	3
99Q2	-19.49	-19.46	-19.61	-19.64	-19.65	-19.58	5
99Q3	-19.86	-19.76	-20.17	-20.13	-20.14	-20.35	6
99Q4	-19.91	-19.80	-20.34	-20.26	-20.27	-20.27	3
00Q1	-19.92	-19.82	-20.29	-20.22	-20.16	-20.08	3
00Q2	-20.09	-20.00	-20.46	-20.38	-20.35	-20.32	3
00Q3	-20.32	-20.24	-20.60	-20.49	-20.43	-20.34	3
00Q4	-20.31	-20.22	-20.54	-20.43	-20.39	-20.28	3
01Q1	-20.36	-20.24	-20.59	-20.50	-20.42	-20.33	3
01Q2	-20.36	-20.29	-20.60	-20.49	-20.39	-20.33	3
01Q3	-20.15	-20.49	-20.42	-20.36	-20.25	-20.18	2
01Q4	-20.23	-20.65	-20.55	-20.44	-20.36	-20.27	2
scores quarter by quarter 02Q1	-20.19	-20.55	-20.46	-20.37	-20.31	-20.20	2
02Q2	-20.14	-20.57	-20.47	-20.37	-20.32	-20.20	2
02Q3	-20.22	-20.63	-20.57	-20.46	-20.47	-20.37	2
02Q4	-20.29	-20.59	-20.65	-20.54	-20.56	-20.48	3
03Q1	-20.40	-20.64	-20.84	-20.73	-20.71	-20.60	3
03Q2	-20.34	-20.64	-20.78	-20.69	-20.62	-20.53	3
03Q3	-20.30	-20.56	-20.65	-20.58	-20.49	-20.38	3
03Q4	-20.31	-20.55	-20.63	-20.60	-20.50	-20.43	3
04Q1	-20.32	-20.51	-20.58	-20.51	-20.41	-20.33	3
04Q2	-20.39	-20.65	-20.80	-20.79	-20.70	-20.60	3
04Q3	-20.40	-20.77	-20.99	-20.97	-20.88	-20.80	3
04Q4	-20.39	-20.66	-20.87	-20.85	-20.82	-20.86	3
05Q1	-20.29	-20.55	-20.82	-20.79	-20.77	-20.81	3
05Q2	-20.39	-20.67	-20.81	-20.81	-20.75	-20.75	4
05Q3	-20.38	-20.68	-20.92	-20.83	-20.76	-20.83	3
average score	**-20.18**	**-20.35**	**-20.53**	**-20.46**	**-20.41**	**-20.35**	**3** **70%**

◊ *Results: one-quarter-ahead forecasts*

factors	One-quarter-ahead forecast: BIC IC_{GDP3} 1	2	3	4	5	6	min
99Q1	-18.90	-18.90	-18.92	-18.84	-18.84	-18.74	3
99Q2	-19.25	-19.23	-19.31	-19.20	-19.33	-19.22	5
99Q3	-19.60	-19.50	-19.83	-19.71	-19.71	-20.12	6
99Q4	-19.65	-19.54	-20.02	-19.92	-19.88	-20.01	3
00Q1	-19.67	-19.55	-19.99	-19.88	-19.81	-19.79	3
00Q2	-19.85	-19.74	-20.16	-20.05	-20.02	-19.94	3
00Q3	-20.12	-20.02	-20.35	-20.25	-20.14	-20.06	3
00Q4	-20.11	-20.01	-20.30	-20.19	-20.10	-20.00	3
01Q1	-20.15	-20.04	-20.36	-20.25	-20.15	-20.05	3
01Q2	-20.16	-20.07	-20.36	-20.26	-20.15	-20.03	3
01Q3	-19.95	-20.23	-20.18	-20.12	-20.01	-19.93	2
01Q4	-19.98	-20.41	-20.31	-20.21	-20.12	-20.02	2
scores quarter by quarter 02Q1	-19.92	-20.29	-20.21	-20.11	-20.08	-19.96	2
02Q2	-19.88	-20.31	-20.21	-20.11	-20.08	-19.96	2
02Q3	-19.91	-20.34	-20.29	-20.18	-20.20	-20.09	2
02Q4	-19.98	-20.28	-20.37	-20.26	-20.20	-20.20	3
03Q1	-20.10	-20.31	-20.54	-20.44	-20.34	-20.35	3
03Q2	-20.07	-20.36	-20.53	-20.42	-20.31	-20.20	3
03Q3	-20.02	-20.31	-20.40	-20.30	-20.21	-20.11	3
03Q4	-20.03	-20.29	-20.37	-20.30	-20.20	-20.09	3
04Q1	-20.04	-20.26	-20.33	-20.24	-20.14	-20.03	3
04Q2	-20.10	-20.36	-20.52	-20.45	-20.34	-20.24	3
04Q3	-20.11	-20.45	-20.65	-20.58	-20.48	-20.40	3
04Q4	-20.13	-20.37	-20.56	-20.48	-20.38	-20.42	3
05Q1	-20.04	-20.27	-20.49	-20.41	-20.30	-20.30	3
05Q2	-20.13	-20.35	-20.46	-20.41	-20.32	-20.30	3
05Q3	-20.13	-20.35	-20.53	-20.46	-20.42	-20.38	3
05Q4	-20.13	-20.35	-20.59	-20.48	-20.46	-20.46	3
average score	**-19.93**	**-20.09**	**-20.26**	**-20.16**	**-20.10**	**-20.05**	**3** **75%**

◊ *Results: two-quarter-ahead forecasts*

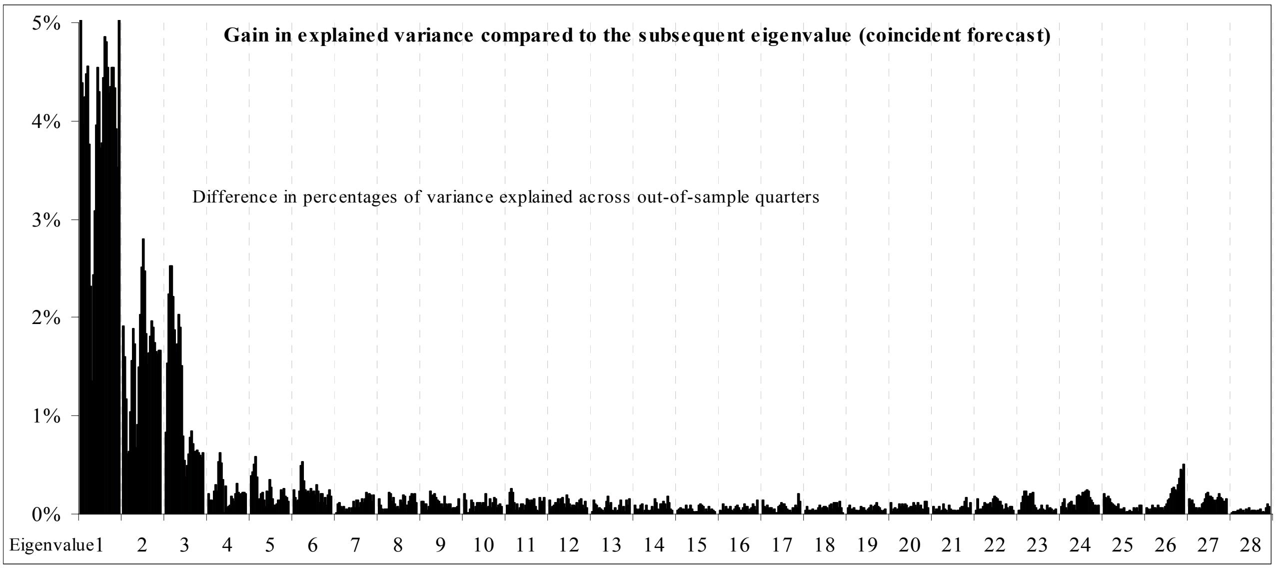
		Two-quarter-ahead forecast: BIC IC_{GDP3}						
factors		1	2	3	4	5	6	min
scores quarter by quarter	99Q2	-18.76	-18.71	-18.75	-18.68	-18.60	-18.79	6
	99Q3	-19.10	-18.98	-19.19	-19.10	-19.00	-19.54	6
	99Q4	-19.25	-19.14	-19.38	-19.31	-19.19	-19.44	6
	00Q1	-19.19	-19.09	-19.40	-19.30	-19.20	-19.14	3
	00Q2	-19.24	-19.15	-19.48	-19.39	-19.30	-19.23	3
	00Q3	-19.85	-19.75	-20.12	-20.02	-19.91	-19.82	3
	00Q4	-19.72	-19.61	-19.90	-19.80	-19.70	-19.61	3
	01Q1	-19.72	-19.61	-19.84	-19.74	-19.63	-19.51	3
	01Q2	-19.77	-19.66	-19.93	-19.82	-19.71	-19.61	3
	01Q3	-19.79	-19.83	-19.98	-19.88	-19.77	-19.69	3
	01Q4	-19.55	-19.60	-19.75	-19.65	-19.54	-19.48	3
	02Q1	-19.53	-19.81	-19.85	-19.75	-19.63	-19.52	3
	02Q2	-19.43	-19.77	-19.73	-19.63	-19.59	-19.48	2
	02Q3	-19.39	-19.77	-19.69	-19.59	-19.54	-19.48	2
	02Q4	-19.44	-19.86	-19.77	-19.70	-19.61	-19.62	2
	03Q1	-19.57	-19.78	-19.89	-19.78	-19.80	-19.70	3
	03Q2	-19.66	-19.70	-20.02	-19.92	-19.81	-19.73	3
	03Q3	-19.60	-19.90	-20.01	-19.90	-19.78	-19.68	3
	03Q4	-19.56	-19.82	-19.92	-19.81	-19.70	-19.59	3
	04Q1	-19.55	-19.82	-19.87	-19.76	-19.65	-19.58	3
	04Q2	-19.57	-19.85	-19.95	-19.84	-19.74	-19.68	3
	04Q3	-19.63	-19.95	-20.10	-20.04	-19.93	-19.87	3
	04Q4	-19.73	-19.96	-20.23	-20.14	-20.02	-19.98	3
	05Q1	-19.64	-19.80	-20.01	-19.89	-19.85	-19.79	3
	05Q2	-19.63	-19.71	-19.93	-19.81	-19.82	-19.74	3
	05Q3	-19.69	-19.73	-19.95	-19.86	-19.81	-19.74	3
	05Q4	-19.69	-19.74	-20.02	-19.94	-19.95	-19.85	3
	06Q1	-19.68	-19.75	-20.06	-20.00	-19.96	-19.86	3
average score		**-19.53**	**-19.64**	**-19.81**	**-19.72**	**-19.63**	**-19.60**	**3** **79%**

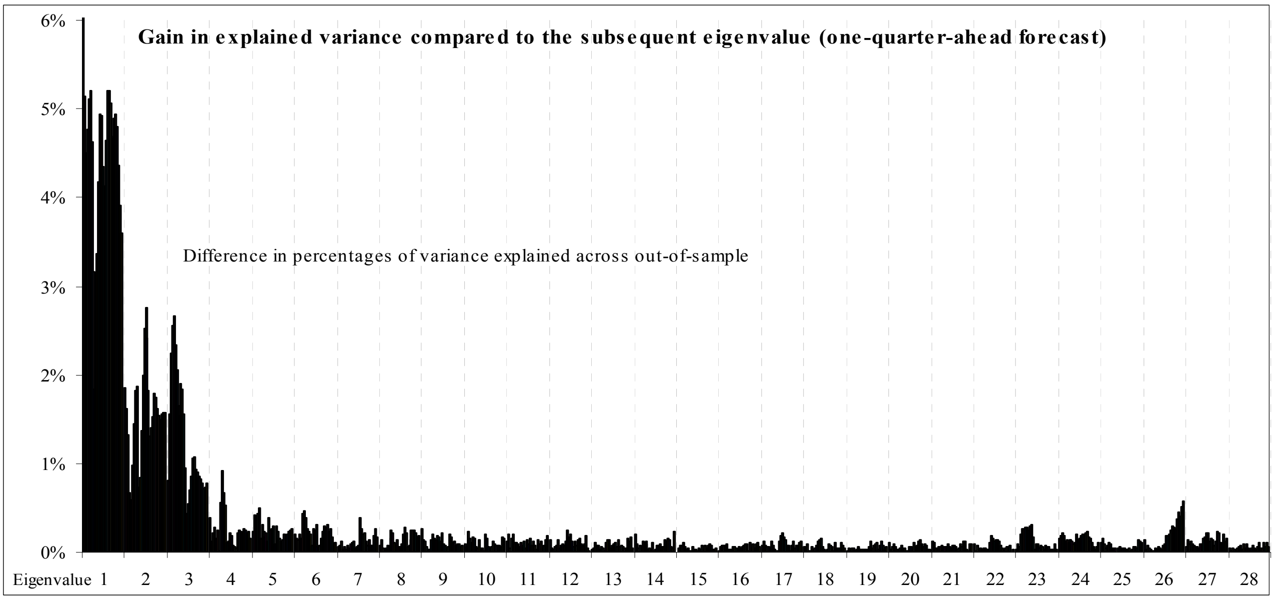
◊ *Results: three-quarter-ahead forecasts*

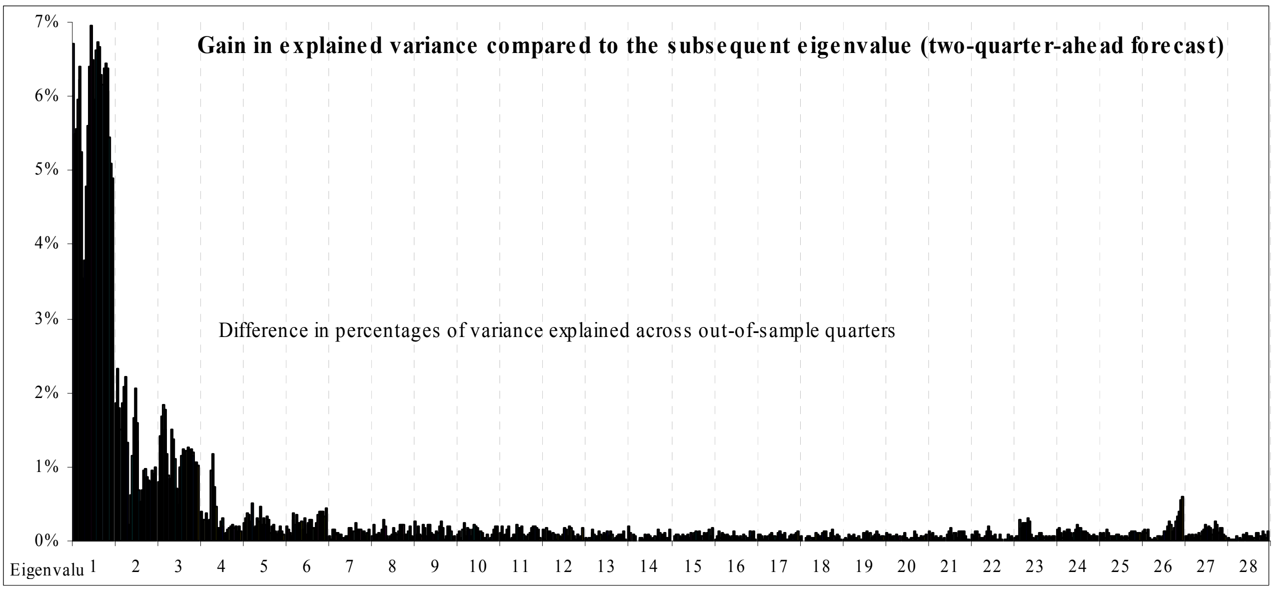
		Three-quarter-ahead forecast: BIC IC_{GDP3}						
factors		1	2	3	4	5	6	min
scores quarter by quarter	99Q3	-18.45	-18.37	-18.51	-18.41	-18.33	-18.33	3
	99Q4	-18.59	-18.49	-18.62	-18.56	-18.51	-18.50	3
	00Q1	-18.64	-18.53	-18.70	-18.62	-18.53	-18.43	3
	00Q2	-18.66	-18.59	-18.75	-18.69	-18.58	-18.48	3
	00Q3	-19.08	-19.00	-19.17	-19.09	-19.00	-18.91	3
	00Q4	-19.34	-19.25	-19.42	-19.32	-19.23	-19.15	3
	01Q1	-19.13	-19.03	-19.22	-19.12	-19.02	-18.91	3
	01Q2	-19.15	-19.05	-19.17	-19.12	-19.01	-18.91	3
	01Q3	-19.22	-19.25	-19.16	-19.24	-19.13	-19.13	2
	01Q4	-19.22	-19.24	-19.24	-19.26	-19.14	-19.09	4
	02Q1	-18.99	-18.97	-19.16	-19.10	-19.03	-19.12	3
	02Q2	-18.98	-19.11	-19.31	-19.23	-19.31	-19.20	5
	02Q3	-18.82	-19.21	-19.15	-19.34	-19.23	-19.19	4
	02Q4	-18.75	-19.23	-19.12	-19.33	-19.22	-19.11	4
	03Q1	-18.86	-19.33	-19.22	-19.34	-19.29	-19.19	4
	03Q2	-19.02	-19.57	-19.47	-19.47	-19.36	-19.25	2
	03Q3	-19.09	-19.14	-19.58	-19.46	-19.36	-19.24	3
	03Q4	-18.99	-19.30	-19.55	-19.44	-19.33	-19.23	3
	04Q1	-18.93	-19.35	-19.45	-19.35	-19.23	-19.12	3
	04Q2	-18.93	-19.42	-19.46	-19.36	-19.25	-19.14	3
	04Q3	-18.95	-19.51	-19.45	-19.42	-19.41	-19.31	2
	04Q4	-19.13	-19.70	-19.62	-19.63	-19.63	-19.56	2
	05Q1	-19.09	-19.71	-19.61	-19.61	-19.57	-19.46	2
	05Q2	-19.09	-19.54	-19.43	-19.43	-19.40	-19.29	2
	05Q3	-19.06	-19.47	-19.36	-19.35	-19.26	-19.15	2
	05Q4	-19.11	-19.34	-19.37	-19.39	-19.28	-19.17	4
	06Q1	-19.11	-19.42	-19.43	-19.42	-19.31	-19.20	3
	06Q2	-19.14	-19.03	-19.84	-19.81	-19.70	-19.72	3
average score		**-18.98**	**-19.18**	**-19.27**	**-19.25**	**-19.17**	**-19.09**	**3** **54%**

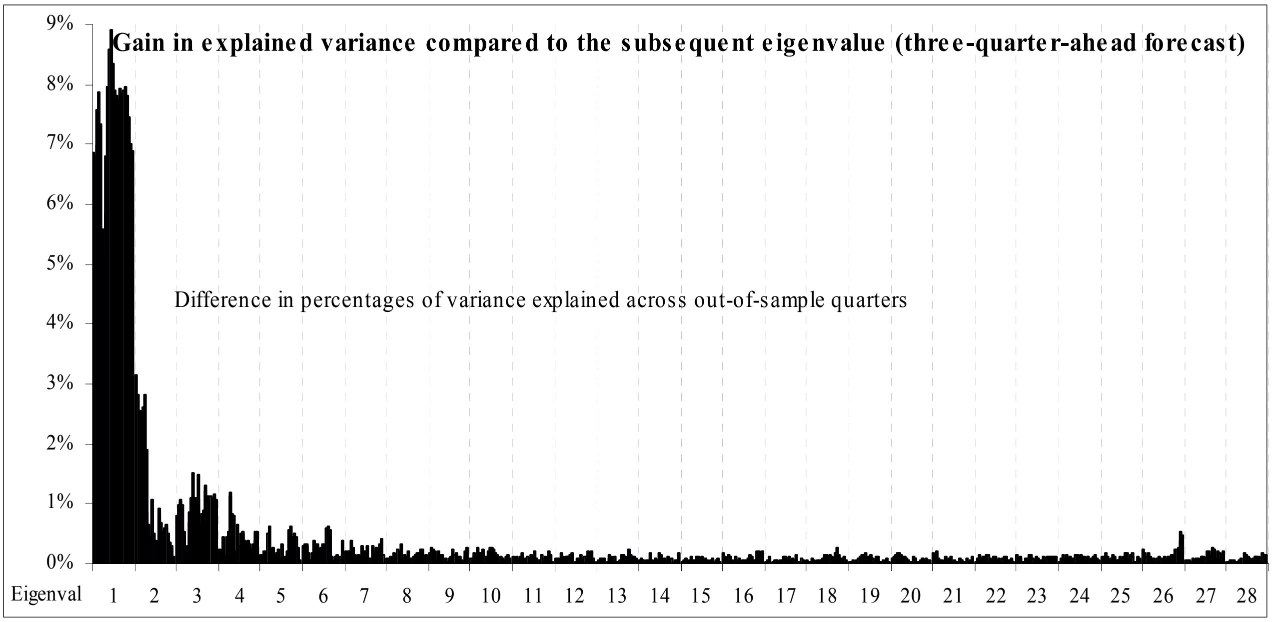
4.3. Eigenvalues "scree" test

◊ *Eigenvalues quarter by quarter*





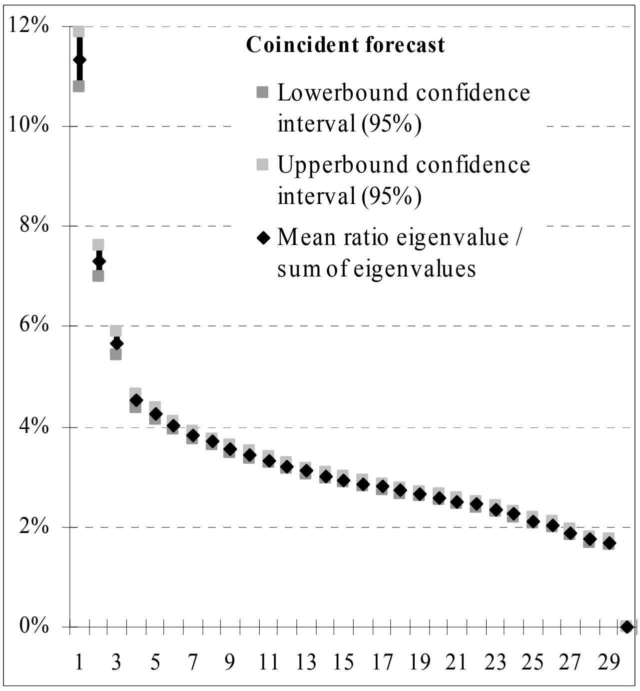


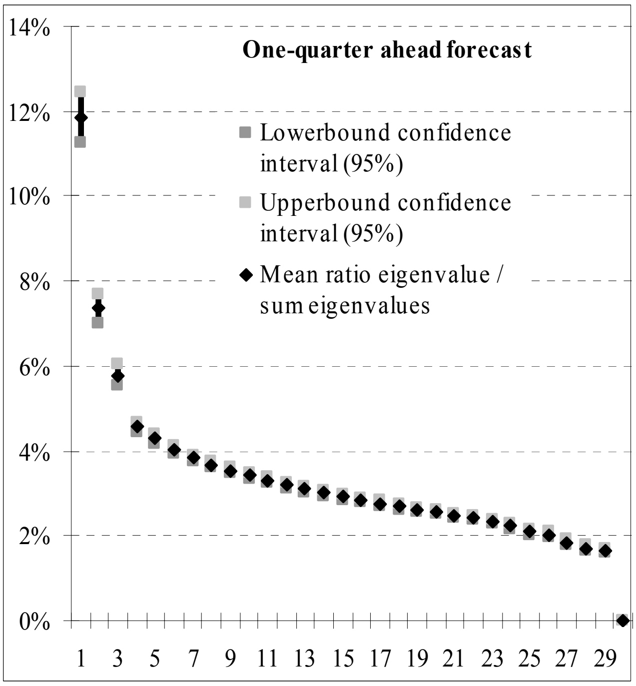


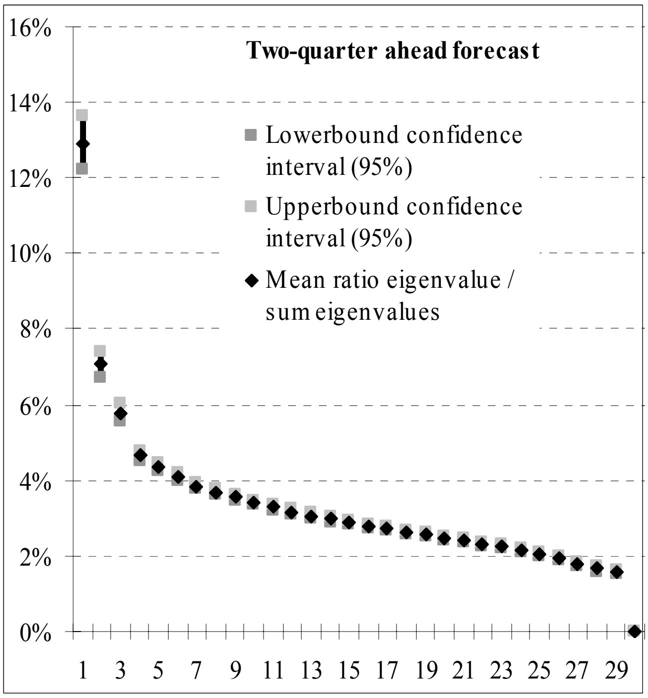
Note to the graphs: eigenvalues estimates are not the same across out-of-sample quarters. For each eigenvalue (on the x-axis), estimates across quarters are stacked between dashed vertical gridlines.

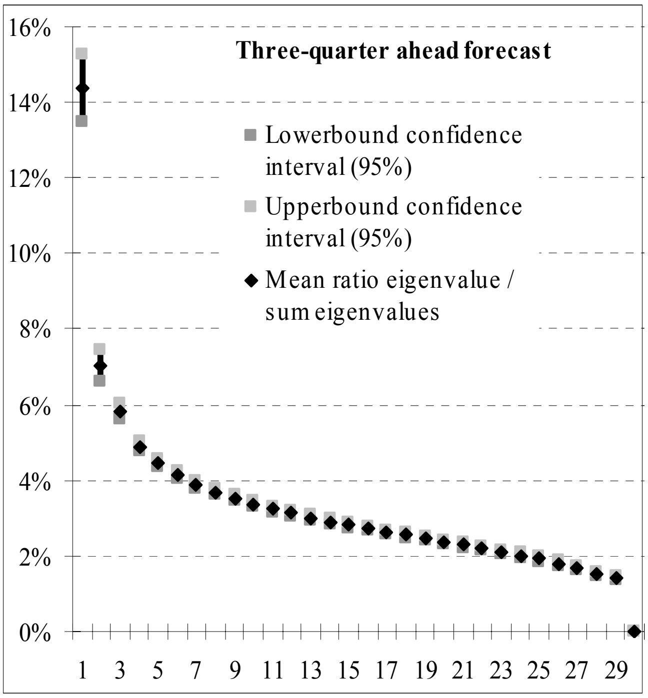
◊ *Bootstrapped eigenvalues*

The following graphs display confidence intervals around eigenvalues estimates (divided by the sum of the eigenvalues) obtained with a cross-sectional bootstrap. Eigenvalues are reported on the x-axis, and percentages of variance on the y-axis.






Two-quarter ahead forecast
Lowerbound confidence interval (95%)
Upperbound confidence interval (95%)
Mean ratio eigenvalue / sum eigenvalues
16%
14%
12%
10%
8%
6%
4%
2%
0%
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29

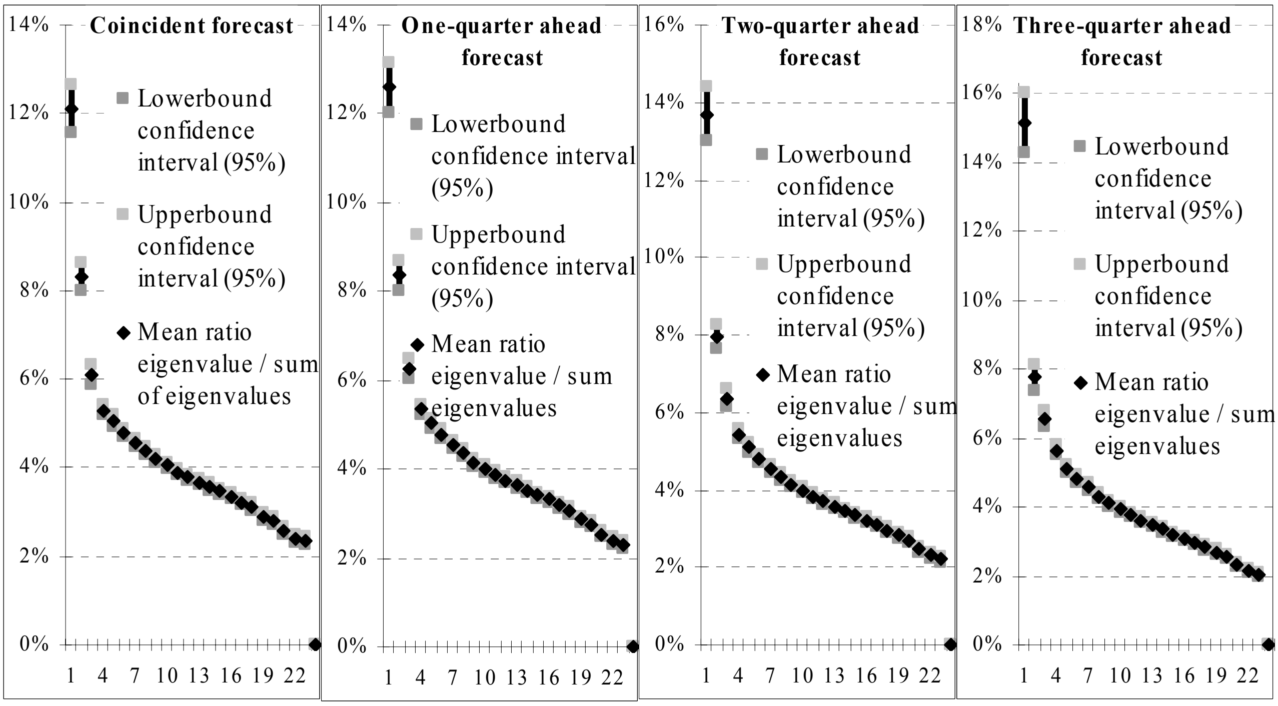

Three-quarter ahead forecast
Lowerbound confidence interval (95%)
Upperbound confidence interval (95%)
Mean ratio eigenvalue / sum eigenvalues
16%
14%
12%
10%
8%
6%
4%
2%
0%
1 3 5 7 9 11 13 15 17 19 21 23 25 27 29

ANNEX 5: SENSITIVITY ANALYSIS

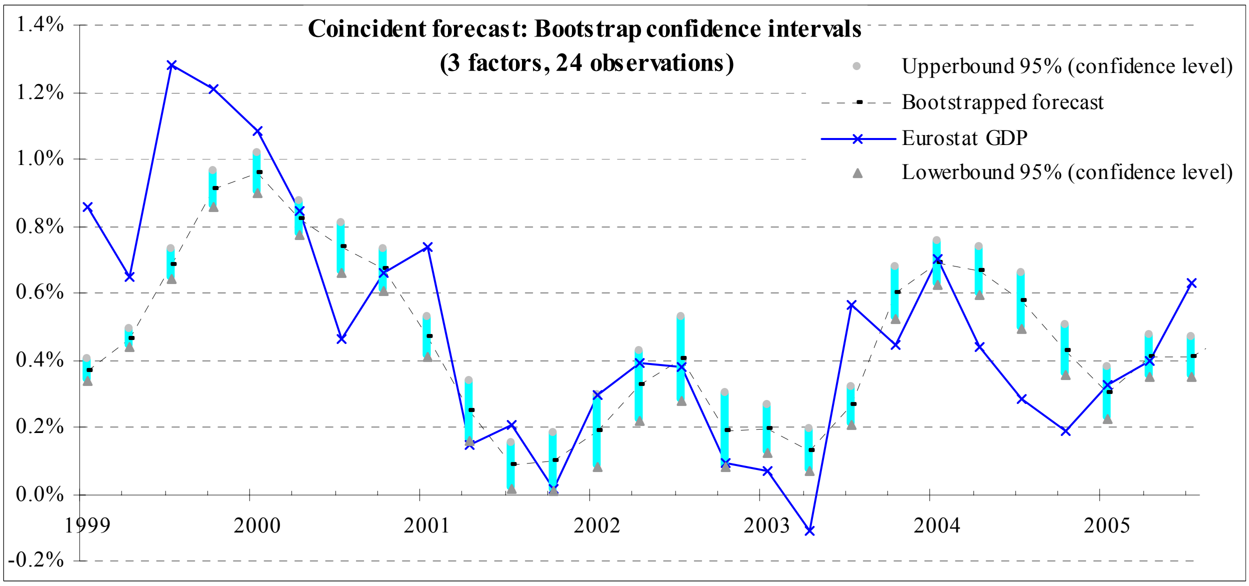
5.1. Time sample span

◊ *Time sample of 24 observations*

Bootstrapped eigenvalues analysis

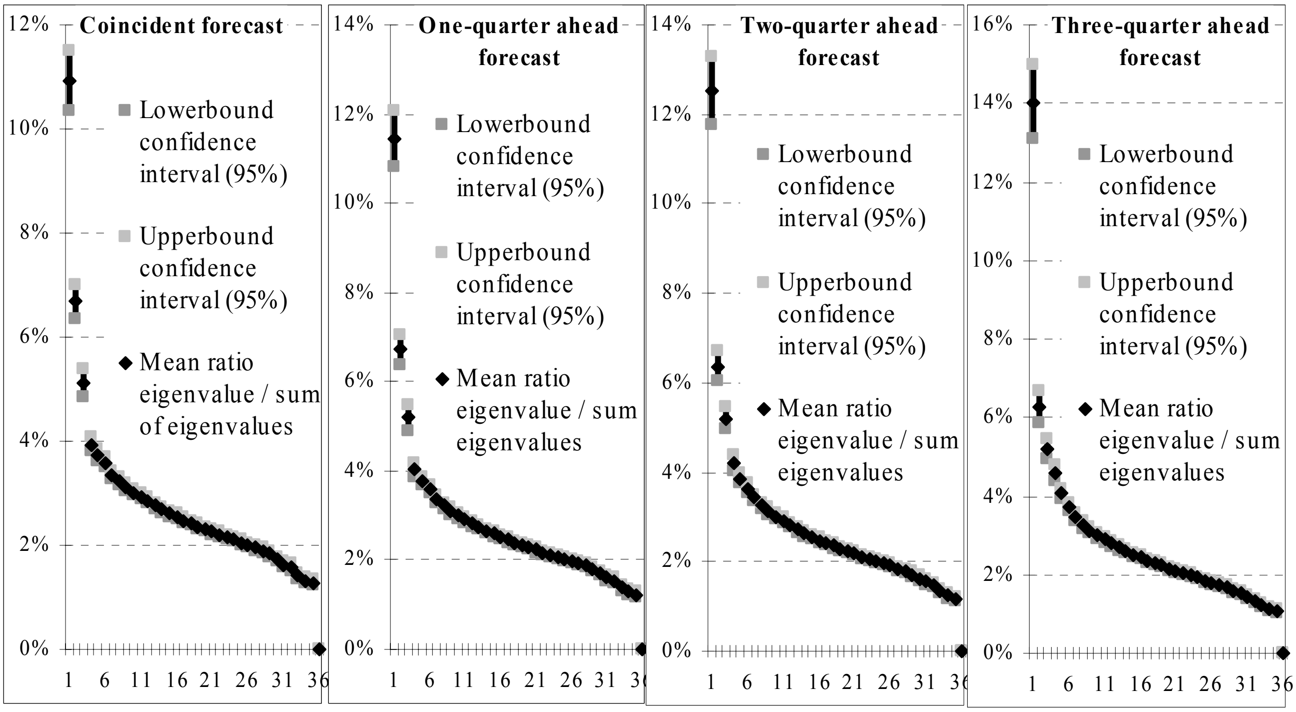


Bootstrapped forecasts confidence intervals

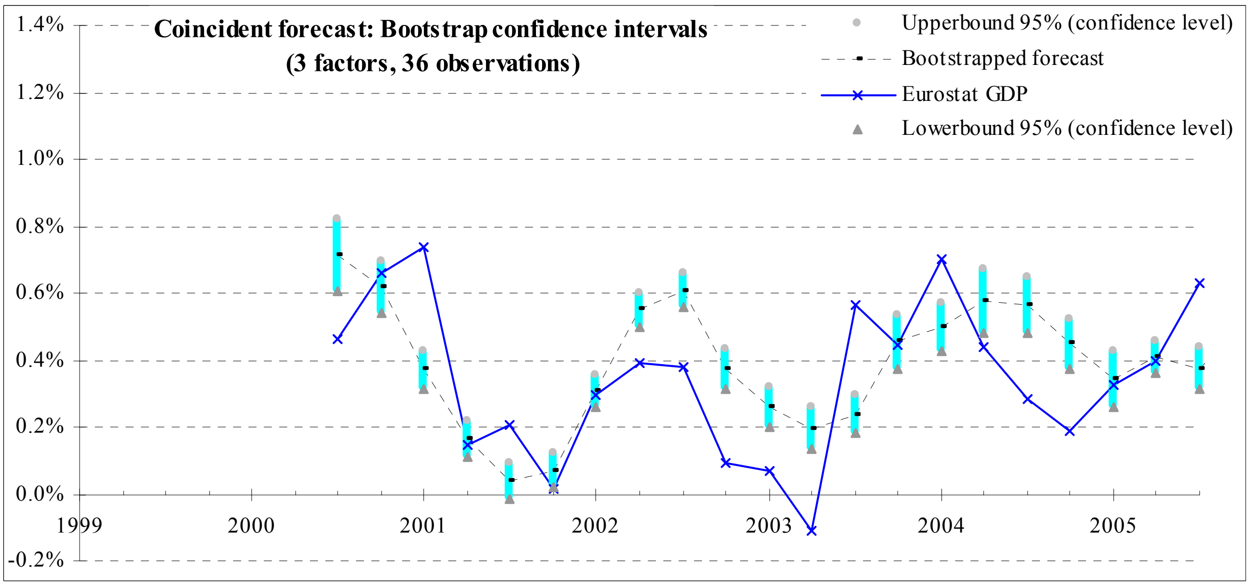


◊ *Time sample of 36 observations*

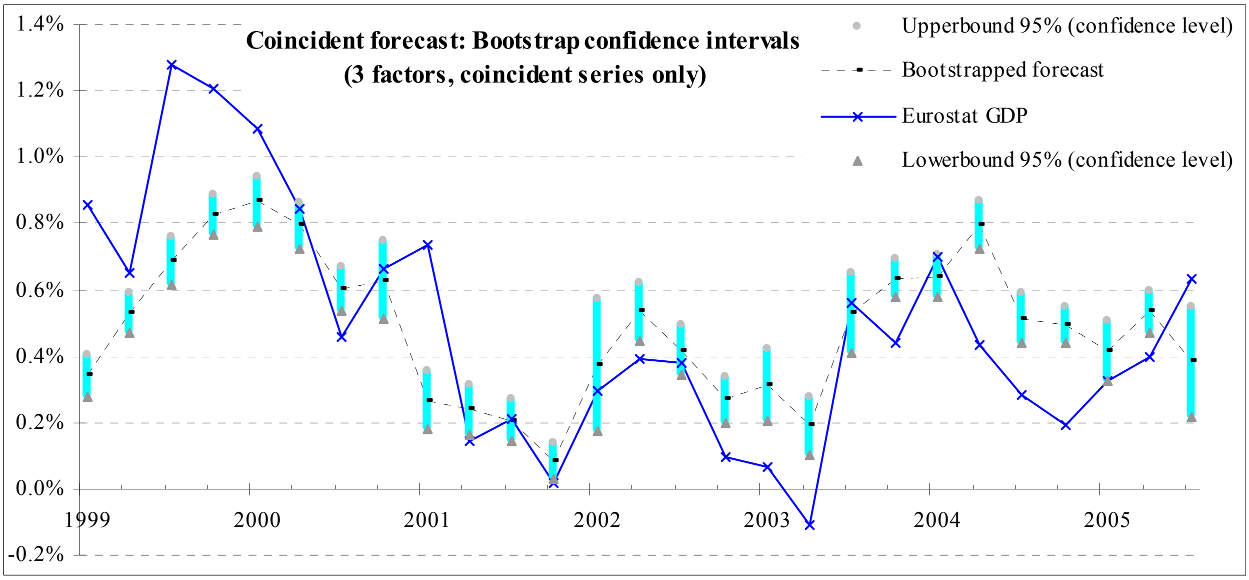
Bootstrapped eigenvalues analysis



Bootstrapped forecasts confidence intervals

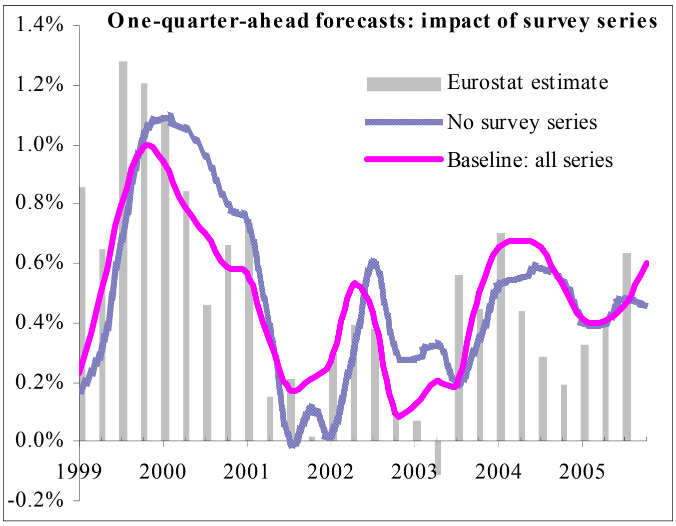


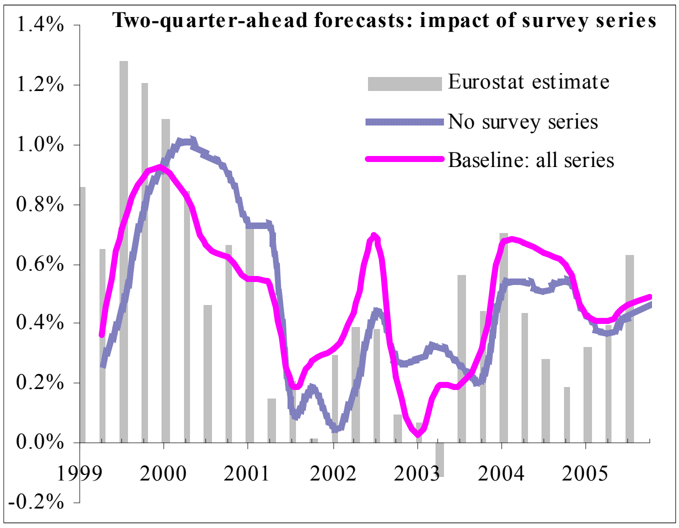
5.2. Number of lags for stacked series

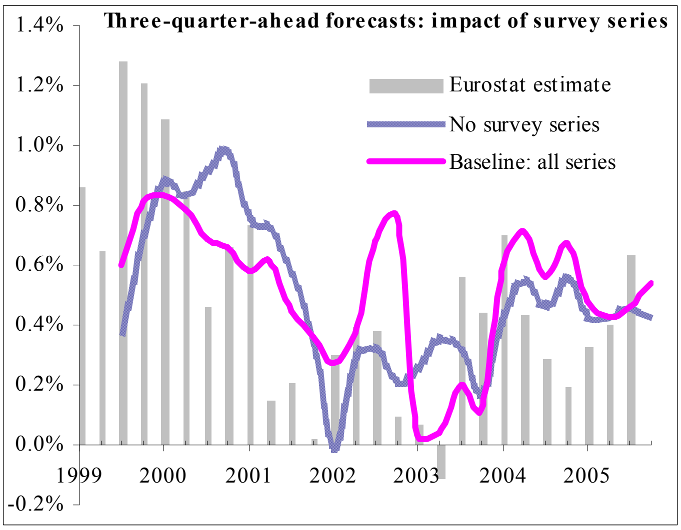


5.3. Data composition

◊ *Compared forecasts with/without survey series*







◊ *Marginal contributions of various subsets of series (three factors used)*

The following tables are not based on bootstrapped simulations (except in the cases of forecasts computed with all series, with all series excluding survey data, and with all series excluding foreign trade data). Since a relatively small subset of data is removed in the cases of non-bootstrapped simulations, bootstrapped forecasts would be very close to sample forecasts.

Note that marginal contributions to the forecasts do not necessarily add up to zero as would be expected with a linear model. For example in 2000Q1, small negative marginal contributions of various subsets of survey data result in a much larger negative contribution of the subset of survey data altogether. This is due to the fact that the estimates of factors are different where a large subset of series is removed. For each subset of data, is displayed the marginal contribution to the forecast

Coincident forecasts

	Marginal contributions of various subsets of data												
Coincident forecast	all survey series (bootstrapped)	consumer confidence	industrial confidence	building confidence	retail confidence	service+financ e confidence	financial variable	foreign trade (bootstrapped)	CPI	PPI	employment	output	sales
mars-99	**0.16%**	0.04%	***0.09%***	0.01%	-0.01%	0.00%	-0.02%	-0.03%	-0.04%	-0.03%	***0.07%***	0.01%	0.00%
juin-99	***0.08%***	0.01%	***0.07%***	0.00%	0.00%	0.00%	0.00%	0.02%	-0.02%	***-0.06%***	-0.01%	0.02%	-0.01%
sept-99	***-0.07%***	-0.03%	0.04%	0.00%	-0.01%	0.00%	0.02%	***0.06%***	0.04%	***-0.06%***	0.00%	0.00%	-0.01%
déc-99	**-0.15%**	-0.04%	0.03%	-0.01%	-0.01%	-0.01%	0.01%	0.04%	0.03%	-0.02%	0.04%	-0.01%	-0.01%
mars-00	**-0.23%**	-0.03%	***-0.06%***	-0.01%	-0.01%	-0.01%	0.00%	***0.06%***	***0.05%***	-0.02%	0.05%	-0.02%	-0.01%
juin-00	**-0.24%**	-0.03%	***-0.10%***	-0.02%	-0.01%	-0.01%	-0.01%	***0.05%***	***0.06%***	-0.02%	***0.07%***	-0.01%	-0.01%
sept-00	**-0.22%**	-0.02%	***-0.09%***	-0.02%	-0.02%	-0.01%	-0.01%	0.04%	***0.05%***	-0.01%	***0.05%***	-0.03%	-0.01%
déc-00	**-0.20%**	-0.01%	**-0.19%**	-0.03%	-0.01%	-0.01%	0.00%	***0.05%***	0.04%	***-0.05%***	***0.06%***	-0.01%	-0.01%
mars-01	**-0.12%**	-0.01%	***-0.08%***	-0.02%	-0.01%	-0.01%	-0.01%	-0.01%	0.04%	-0.03%	***0.05%***	0.02%	0.00%
juin-01	***0.07%***	-0.01%	**0.13%**	-0.01%	-0.01%	0.00%	-0.01%	-0.04%	0.01%	0.01%	0.01%	-0.01%	0.00%
sept-01	**0.14%**	-0.02%	**0.17%**	0.00%	-0.01%	0.00%	0.00%	***-0.06%***	-0.02%	0.02%	0.01%	0.00%	0.00%
déc-01	**0.12%**	-0.03%	**0.15%**	0.00%	-0.01%	0.01%	-0.01%	-0.02%	-0.02%	0.02%	-0.02%	-0.01%	0.01%
mars-02	**0.19%**	-0.01%	**0.20%**	-0.01%	0.00%	0.01%	-0.01%	-0.03%	***-0.06%***	-0.01%	**-0.12%**	0.05%	0.00%
juin-02	**0.17%**	0.03%	**0.16%**	-0.03%	-0.01%	0.00%	-0.02%	-0.04%	**0.10%**	0.03%	**0.14%**	-0.01%	0.00%
sept-02	**-0.13%**	**-0.13%**	**-0.14%**	0.01%	***-0.05%***	-0.03%	***-0.05%***	**-0.21%**	**0.17%**	**0.22%**	***0.10%***	-0.01%	0.03%
déc-02	**-0.20%**	**-0.14%**	***-0.09%***	0.02%	-0.03%	-0.01%	0.00%	**-0.15%**	***0.06%***	**0.15%**	0.01%	0.01%	0.00%
mars-03	**-0.16%**	**-0.12%**	-0.04%	0.00%	-0.01%	0.01%	-0.02%	-0.02%	***0.06%***	0.04%	-0.01%	0.00%	0.00%
juin-03	***-0.08%***	***-0.07%***	-0.03%	0.01%	-0.01%	0.01%	-0.02%	-0.04%	**0.11%**	0.04%	0.01%	-0.01%	-0.01%
sept-03	0.01%	-0.03%	-0.01%	0.01%	-0.01%	0.03%	0.00%	0.02%	0.03%	-0.01%	-0.05%	0.01%	-0.01%
déc-03	**0.12%**	0.01%	0.04%	0.01%	0.01%	0.03%	0.02%	***0.05%***	**-0.15%**	***-0.08%***	***-0.09%***	0.00%	0.00%
mars-04	**0.12%**	0.04%	-0.01%	0.02%	0.01%	0.03%	0.00%	-0.01%	***0.09%***	0.04%	-0.03%	-0.01%	0.00%
juin-04	***0.09%***	0.04%	-0.02%	0.01%	0.00%	0.01%	-0.02%	-0.04%	0.04%	0.04%	-0.04%	0.00%	0.00%
sept-04	0.04%	0.03%	-0.05%	0.02%	0.00%	0.02%	-0.01%	0.00%	-0.02%	0.00%	-0.04%	-0.02%	0.00%
déc-04	0.01%	***0.05%***	***-0.08%***	0.02%	0.00%	0.01%	-0.01%	-0.01%	***0.08%***	0.02%	-0.02%	-0.03%	-0.01%
mars-05	0.01%	***0.06%***	***-0.09%***	0.02%	0.00%	0.02%	0.00%	-0.03%	0.04%	0.00%	-0.04%	-0.03%	0.00%
juin-05	-0.04%	0.04%	***-0.09%***	0.01%	0.00%	0.01%	0.01%	0.01%	0.01%	-0.04%	-0.03%	0.00%	0.00%
sept-05	0.01%	0.03%	-0.05%	0.01%	0.00%	0.01%	0.00%	-0.03%	0.02%	-0.05%	-0.02%	0.01%	0.00%

One-quarter ahead forecasts

	Marginal contributions of various subsets of data												
1Q ahead forecast	all survey series (bootstrapped)	consumer confidence	industrial confidence	building confidence	retail confidence	service+financ e confidence	financial variable	foreign trade (bootstrapped)	CPI	PPI	employment	output	sales
mars-99	***0.06%***	0.04%	***0.05%***	0.04%	-0.02%	-0.01%	-0.02%	***-0.07%***	-0.02%	0.03%	***0.06%***	-0.01%	-0.01%
juin-99	**0.20%**	0.02%	**0.13%**	0.02%	0.00%	0.00%	0.00%	-0.01%	0.00%	-0.02%	0.00%	0.00%	-0.02%
sept-99	**0.11%**	-0.02%	**0.11%**	-0.02%	0.00%	0.00%	0.02%	0.03%	***0.05%***	***-0.07%***	-0.01%	0.00%	-0.02%
déc-99	-0.04%	-0.04%	***0.07%***	-0.04%	-0.01%	0.00%	0.03%	0.04%	0.05%	-0.03%	0.03%	-0.02%	-0.01%
mars-00	**-0.15%**	-0.03%	-0.02%	-0.03%	-0.01%	-0.01%	0.00%	0.03%	***0.07%***	-0.03%	0.04%	-0.01%	-0.01%
juin-00	**-0.26%**	-0.03%	**-0.11%**	-0.03%	-0.01%	-0.01%	-0.01%	0.04%	***0.09%***	-0.03%	***0.06%***	-0.01%	-0.01%
sept-00	**-0.25%**	-0.03%	***-0.10%***	-0.03%	-0.02%	-0.01%	-0.01%	0.03%	***0.07%***	-0.02%	0.05%	-0.01%	0.00%
déc-00	**-0.20%**	-0.02%	**-0.21%**	-0.02%	-0.02%	0.00%	0.02%	0.02%	***0.06%***	-0.04%	0.05%	-0.01%	-0.01%
mars-01	**-0.17%**	0.00%	**-0.21%**	0.00%	-0.01%	-0.01%	0.01%	0.04%	***0.05%***	***-0.05%***	0.05%	0.01%	0.00%
juin-01	-0.05%	-0.02%	***0.07%***	-0.02%	-0.01%	-0.01%	-0.01%	-0.04%	0.02%	-0.01%	0.01%	0.03%	0.01%
sept-01	**0.18%**	-0.03%	**0.18%**	-0.03%	-0.01%	0.00%	0.00%	-0.04%	-0.03%	0.01%	0.00%	-0.01%	0.01%
déc-01	**0.10%**	-0.03%	**0.13%**	-0.03%	-0.01%	0.00%	-0.03%	0.00%	-0.03%	0.01%	-0.02%	0.02%	0.01%
mars-02	**0.25%**	-0.02%	**0.21%**	-0.02%	0.00%	0.01%	0.00%	-0.01%	***-0.06%***	-0.01%	0.00%	0.00%	0.00%
juin-02	**0.22%**	0.02%	**0.25%**	0.02%	-0.04%	-0.02%	-0.03%	***-0.06%***	**0.12%**	-0.02%	**0.17%**	-0.02%	-0.04%
sept-02	**-0.17%**	**-0.11%**	***-0.08%***	**-0.11%**	-0.03%	0.02%	-0.03%	**-0.17%**	**0.18%**	**0.28%**	**0.11%**	0.03%	***0.05%***
déc-02	**-0.20%**	**-0.15%**	***-0.08%***	**-0.15%**	-0.03%	0.00%	-0.01%	**-0.16%**	**0.14%**	**0.26%**	0.04%	0.03%	0.02%
mars-03	**-0.15%**	**-0.13%**	-0.01%	**-0.13%**	-0.01%	0.00%	-0.02%	-0.03%	***0.09%***	***0.08%***	-0.01%	0.02%	0.02%
juin-03	**-0.12%**	***-0.08%***	-0.03%	***-0.08%***	-0.01%	0.01%	-0.02%	0.00%	***0.07%***	0.01%	-0.01%	0.01%	0.00%
sept-03	0.00%	-0.03%	-0.01%	-0.03%	-0.01%	0.02%	-0.02%	-0.02%	**0.12%**	0.00%	-0.03%	-0.01%	-0.01%
déc-03	**0.16%**	0.01%	***0.08%***	0.01%	0.01%	0.03%	0.02%	***0.06%***	**-0.15%**	***-0.06%***	***-0.08%***	0.00%	0.00%
mars-04	**0.13%**	0.04%	0.00%	0.04%	0.01%	0.02%	0.01%	0.01%	0.01%	-0.01%	***-0.06%***	-0.01%	-0.01%
juin-04	**0.12%**	0.05%	-0.01%	0.05%	0.00%	0.01%	-0.02%	-0.03%	***0.07%***	0.03%	-0.05%	-0.01%	-0.01%
sept-04	***0.07%***	0.02%	0.00%	0.02%	0.00%	0.01%	-0.02%	0.00%	-0.04%	0.00%	-0.04%	0.00%	0.01%
déc-04	-0.01%	***0.05%***	***-0.07%***	***0.05%***	0.00%	0.00%	-0.02%	0.01%	***0.07%***	-0.02%	-0.03%	-0.01%	-0.01%
mars-05	0.01%	***0.05%***	***-0.08%***	***0.05%***	-0.01%	0.02%	0.00%	-0.01%	0.04%	-0.03%	-0.02%	-0.01%	0.00%
juin-05	0.00%	0.05%	***-0.07%***	0.05%	0.00%	0.01%	0.01%	-0.01%	0.03%	-0.04%	-0.01%	-0.01%	0.00%
sept-05	-0.03%	0.01%	***-0.06%***	0.01%	0.00%	0.00%	0.01%	0.02%	0.02%	***-0.05%***	-0.01%	0.00%	0.00%
déc-05	**0.14%**	-0.01%	***0.05%***	-0.01%	0.01%	0.00%	0.00%	-0.01%	0.03%	***-0.05%***	-0.01%	0.00%	0.00%

Two-quarter ahead forecasts

Marginal contributions of various subsets of data													
2Q ahead forecast	all survey series (bootstrapped)	consumer confidence	industrial confidence	building confidence	retail confidence	service+financ e confidence	financial variable	foreign trade (bootstrapped)	CPI	PPI	employment	output	sales
juin-99	**0.11%**	***0.07%***	0.05%	***0.07%***	-0.02%	0.00%	0.02%	-0.05%	0.04%	0.04%	0.02%	0.00%	-0.02%
sept-99	**0.27%**	0.03%	**0.16%**	0.03%	0.00%	0.00%	0.05%	0.02%	**0.10%**	***-0.08%***	-0.03%	-0.01%	-0.03%
déc-99	**0.10%**	-0.03%	**0.14%**	-0.03%	-0.01%	0.00%	***0.07%***	0.03%	***0.09%***	***-0.07%***	0.00%	-0.02%	-0.03%
mars-00	-0.01%	-0.04%	**0.15%**	-0.04%	-0.01%	0.00%	0.03%	0.05%	**0.10%**	-0.04%	0.03%	-0.02%	-0.03%
juin-00	**-0.18%**	-0.01%	**-0.12%**	-0.01%	-0.01%	-0.01%	-0.02%	0.04%	**0.12%**	***-0.06%***	***0.05%***	0.00%	-0.02%
sept-00	**-0.31%**	-0.03%	**-0.19%**	-0.03%	-0.02%	0.00%	0.00%	***0.06%***	**0.10%**	-0.05%	0.04%	-0.01%	-0.01%
déc-00	**-0.28%**	-0.01%	**-0.26%**	-0.01%	-0.03%	0.00%	0.01%	0.03%	***0.10%***	***-0.09%***	0.04%	-0.01%	-0.03%
mars-01	**-0.18%**	-0.01%	**-0.23%**	-0.01%	-0.02%	0.00%	0.03%	0.03%	***0.09%***	**-0.14%**	***0.06%***	0.00%	-0.02%
juin-01	**-0.19%**	0.00%	**-0.14%**	0.00%	-0.02%	-0.01%	0.02%	0.04%	***0.06%***	**-0.15%**	0.03%	0.03%	-0.01%
sept-01	***0.09%***	***-0.05%***	-0.04%	***-0.05%***	-0.02%	0.00%	-0.02%	***-0.09%***	-0.02%	0.00%	0.00%	0.03%	0.02%
déc-01	***0.09%***	-0.05%	**0.12%**	-0.05%	-0.01%	0.01%	-0.01%	0.02%	***-0.07%***	0.00%	-0.02%	-0.01%	0.01%
mars-02	**0.27%**	***-0.05%***	**0.19%**	***-0.05%***	0.01%	0.00%	-0.04%	-0.01%	***-0.07%***	-0.02%	-0.02%	0.05%	0.00%
juin-02	**0.27%**	-0.01%	**0.21%**	-0.01%	0.01%	0.01%	-0.01%	-0.01%	***-0.06%***	0.00%	0.00%	-0.01%	0.00%
sept-02	**0.22%**	***-0.09%***	**0.23%**	***-0.09%***	-0.02%	0.01%	**-0.12%**	**-0.15%**	***-0.09%***	0.03%	0.04%	***0.06%***	0.02%
déc-02	***-0.08%***	**-0.20%**	-0.03%	**-0.20%**	-0.04%	0.01%	-0.03%	**-0.23%**	**0.17%**	**0.39%**	***0.06%***	-0.01%	0.03%
mars-03	**-0.26%**	**-0.24%**	***-0.07%***	**-0.24%**	-0.05%	0.00%	***-0.06%***	***-0.07%***	**0.10%**	**0.24%**	-0.01%	0.04%	0.04%
juin-03	**-0.15%**	**-0.12%**	0.00%	**-0.12%**	-0.02%	0.00%	-0.01%	0.00%	0.03%	0.03%	-0.04%	0.01%	0.01%
sept-03	***-0.06%***	***-0.07%***	-0.01%	***-0.07%***	-0.01%	0.01%	-0.02%	-0.01%	***0.06%***	-0.01%	-0.02%	-0.01%	0.00%
déc-03	**0.10%**	-0.03%	0.05%	-0.03%	0.01%	0.02%	-0.03%	0.04%	-0.01%	-0.03%	***-0.07%***	0.00%	0.00%
mars-04	**0.15%**	0.04%	0.04%	0.04%	0.01%	0.03%	0.02%	0.04%	***-0.09%***	-0.03%	***-0.07%***	0.01%	0.00%
juin-04	**0.13%**	***0.05%***	0.00%	***0.05%***	0.00%	0.02%	0.01%	-0.02%	0.05%	0.01%	***-0.05%***	-0.01%	-0.01%
sept-04	**0.13%**	0.03%	-0.01%	0.03%	0.01%	0.02%	-0.03%	0.00%	0.03%	0.02%	-0.04%	-0.01%	0.01%
déc-04	***0.05%***	0.04%	-0.01%	0.04%	0.00%	0.00%	-0.03%	0.02%	-0.01%	-0.03%	-0.03%	0.00%	0.00%
mars-05	0.01%	***0.08%***	***-0.07%***	***0.08%***	0.00%	0.01%	-0.01%	0.00%	***0.06%***	-0.03%	-0.02%	-0.01%	-0.01%
juin-05	0.04%	0.04%	-0.05%	0.04%	0.00%	0.02%	-0.01%	0.01%	0.01%	-0.04%	-0.02%	0.00%	0.01%
sept-05	0.04%	***0.05%***	-0.04%	***0.05%***	0.00%	0.00%	0.01%	0.01%	0.04%	-0.05%	-0.01%	-0.01%	0.00%
déc-05	0.02%	0.00%	-0.01%	0.00%	0.00%	0.00%	0.00%	0.01%	0.04%	-0.05%	-0.01%	0.00%	-0.01%
mars-06	**0.20%**	-0.04%	***0.08%***	0.01%	0.01%	0.01%	0.01%	-0.01%	-0.02%	-0.04%	0.00%	0.00%	0.00%

Three-quarter ahead forecasts

	Marginal contributions of various subsets of data												
3Q ahead forecast	all survey series (bootstrapped)	consumer confidence	industrial confidence	building confidence	retail confidence	service+financ e confidence	financial variable	foreign trade (bootstrapped)	CPI	PPI	employment	output	sales
sept-99	**0.25%**	**0.10%**	***0.06%***	**0.10%**	-0.01%	0.00%	0.04%	-0.01%	**0.14%**	0.00%	-0.03%	-0.01%	-0.02%
déc-99	**0.14%**	0.01%	**0.16%**	0.01%	-0.01%	0.00%	**0.14%**	0.02%	**0.15%**	***-0.05%***	0.00%	-0.02%	-0.04%
mars-00	-0.03%	-0.03%	**0.12%**	-0.03%	-0.03%	0.00%	***0.10%***	0.05%	**0.13%**	***-0.06%***	0.01%	-0.02%	-0.05%
juin-00	-0.04%	0.00%	***-0.06%***	0.00%	0.01%	0.00%	0.00%	***0.06%***	**0.18%**	***-0.09%***	***0.06%***	-0.01%	-0.05%
sept-00	**-0.22%**	-0.01%	**-0.14%**	-0.01%	-0.03%	0.00%	-0.02%	0.04%	**0.13%**	-0.04%	0.04%	0.01%	0.00%
déc-00	**-0.30%**	0.03%	***-0.09%***	0.03%	-0.03%	0.00%	**0.10%**	**0.13%**	**0.19%**	**-0.12%**	***0.06%***	-0.03%	***-0.07%***
mars-01	**-0.17%**	0.03%	**-0.16%**	0.03%	-0.03%	0.00%	**0.10%**	***0.07%***	**0.19%**	**-0.13%**	***0.07%***	-0.03%	***-0.09%***
juin-01	-0.05%	***0.09%***	-0.01%	***0.09%***	-0.04%	0.01%	**0.14%**	**0.17%**	**0.21%**	0.00%	-0.02%	0.01%	0.00%
sept-01	**-0.14%**	0.05%	**-0.12%**	0.05%	-0.02%	0.00%	0.00%	***0.06%***	***0.09%***	**-0.19%**	0.02%	0.00%	-0.04%
déc-01	0.04%	0.01%	**-0.16%**	0.01%	-0.04%	0.00%	***-0.10%***	***0.06%***	0.04%	**-0.14%**	0.02%	0.02%	0.03%
mars-02	**0.28%**	***-0.08%***	***0.10%***	***-0.08%***	0.01%	0.00%	-0.04%	0.00%	**-0.16%**	***0.07%***	***-0.05%***	0.01%	0.02%
juin-02	***0.08%***	***-0.10%***	**0.10%**	***-0.10%***	0.00%	0.00%	***-0.07%***	0.00%	***-0.06%***	**0.10%**	-0.03%	0.04%	0.02%
sept-02	**0.36%**	***0.06%***	**0.23%**	***0.06%***	0.01%	0.01%	0.01%	0.01%	***-0.07%***	***-0.07%***	-0.04%	-0.01%	-0.01%
déc-02	**0.53%**	**-0.10%**	**0.42%**	**-0.10%**	0.00%	-0.01%	**-0.13%**	***0.08%***	***0.06%***	***0.10%***	***0.08%***	-0.02%	-0.03%
mars-03	**-0.23%**	**-0.30%**	-0.02%	**-0.30%**	-0.04%	0.00%	***-0.06%***	**-0.10%**	***0.08%***	0.34%	-0.02%	0.04%	***0.06%***
juin-03	**-0.32%**	**-0.24%**	***-0.10%***	**-0.24%**	***-0.07%***	0.00%	***-0.10%***	-0.01%	**0.18%**	**0.19%**	-0.03%	***0.05%***	0.04%
sept-03	**-0.12%**	**-0.11%**	0.01%	**-0.11%**	-0.02%	0.01%	0.01%	-0.01%	0.01%	0.01%	-0.04%	0.00%	0.01%
déc-03	***-0.06%***	***-0.09%***	***-0.06%***	***-0.09%***	0.00%	0.01%	-0.03%	-0.02%	**0.10%**	-0.02%	-0.04%	-0.01%	0.00%
mars-04	**0.17%**	0.03%	0.00%	0.03%	0.03%	0.03%	***-0.05%***	***0.05%***	***-0.07%***	0.00%	***-0.08%***	0.01%	0.02%
juin-04	**0.17%**	***0.06%***	***0.10%***	***0.06%***	0.00%	0.01%	***0.05%***	-0.01%	***-0.06%***	-0.02%	-0.05%	0.00%	0.00%
sept-04	**0.10%**	0.02%	-0.01%	0.02%	0.00%	0.01%	-0.03%	-0.02%	0.03%	0.04%	-0.05%	0.00%	0.01%
déc-04	**0.12%**	0.02%	0.00%	0.02%	0.01%	0.01%	-0.04%	0.01%	0.02%	-0.01%	***-0.05%***	0.00%	0.01%
mars-05	0.05%	***0.08%***	-0.01%	***0.08%***	-0.01%	0.00%	0.00%	0.00%	0.03%	-0.05%	-0.03%	0.00%	-0.01%
juin-05	0.00%	***0.06%***	***-0.05%***	***0.06%***	0.01%	0.00%	-0.02%	0.00%	0.02%	-0.02%	-0.02%	0.00%	0.01%
sept-05	0.01%	0.03%	-0.05%	0.03%	-0.01%	0.00%	-0.01%	0.02%	***0.06%***	-0.04%	-0.02%	0.00%	0.00%
déc-05	**0.12%**	***0.06%***	0.01%	***0.06%***	0.01%	0.00%	0.00%	0.02%	0.03%	***-0.06%***	0.00%	-0.03%	-0.01%
mars-06	***0.05%***	-0.04%	0.04%	0.00%	0.00%	0.00%	0.02%	-0.03%	0.01%	-0.02%	0.00%	-0.01%	-0.01%
juin-06	**0.15%**	**-0.12%**	0.02%	***0.05%***	0.01%	0.02%	0.04%	0.03%	***0.09%***	-0.04%	0.00%	0.01%	0.01%